Considerations for Modelling Realistic Human Behaviour in ABSS: In pursuit of an ill-defined pipe dream

Gabriella Low Chew Tung[1], Joppe Roosendaal[*], Yashi Punia[*]

[1,*] Delft University of Technology, Faculty of Technology Policy and Management 2628 BX, The Netherlands
[*] indicates that all authors have contributed equally

Abstract
To model and simulate social phenomena accurately requires that human behaviour be captured in the model at a sufficiently high resolution such that realistic complex behaviour can be generated. Despite this, in the field of agent-based social simulations(ABSS), there is a tendency to use simple rules to guide agents' decision-making processes leading to rudimentary and unrealistic agent behaviour. Although there exist several computational frameworks that claim to address this issue, they remain underutilised in social simulations. In this research paper, three computational frameworks and their inner workings were studied and evaluated through literature review for the purpose of modelling human behaviour. This work then compared the usefulness of these frameworks against a set of criteria and identified their capabilities, limitations and the challenges hindering the popularisation of these frameworks.

Keywords: Human Behaviour Modelling, Social Simulations, Agent Based Modelling, BDI, PECS, NoA, Agent Architecture

1 Introduction

Agent-based social simulation (ABSS) is an interdisciplinary field of study that integrates the research areas of computer simulation, social sciences and agent-based computing to produce agent-based models (ABMs) for the simulation of social phenomena (Davidsson, 2002). The field initially emerged in the 1970s with early influential works such as Schelling's Segregation Model and later Epstein and Axtel's Sugarscape model. The ABSS field has continued to develop with its recent rapid expansion attributed to advancements in other disciplines such as artificial intelligence. These advancements have the potential to facilitate more accurate modelling and simulation research in many ABSS application fields including: evacuation, social interaction, economic studies, psychology and social studies (Wijermans, 2023). Consequently, this accelerated development has led to confusion among ABSS scientists, who are working in many different fields, as to which of the many proposed methods or techniques should be implemented for their particular application (Li et al., 2008).

An ongoing debate in ABSS, concerns the best approach to modelling human behaviour (Balke & Gilbert, 2014). Human behaviour modelling can be highly complex. Humans are complex beings who are capable of reactive as well as deliberate behaviour. This can be influenced by numerous factors such as personality, norms, emotions, and past experiences. A model on human behaviour needs to be able to reflect these abilities. The choice of which parts are necessary for a model is up to the modeller. However, the high level of interconnectivity between all of these factors can be a big challenge for any modeller (Goldspink, 2000).

Traditionally, ABSS researchers have sought to implement agent behaviour using simple rules to preserve comprehensibility and explainability of their outcomes i.e. emergent behaviour. This method is referred to as the production-rule system and is one of the most popular approaches to designing agent behaviour. However, it is often criticised for its lack of realism in representing human behaviour (Balke & Gilbert, 2014). This simplicity first approach is captured by the 'Keep It Simple Stupid' (KISS) paradigm as it is referred to in Edmonds and Moss (2004). They argued against the de facto standard of assuming simpler models lead to better or more accurate outcomes. Instead, they advocate for the use of more descriptive models that are able to utilise all available data which are only simplified if needed – 'Keep It Descriptive Stupid' (KIDS). Adam and Gaudou (2016) relate this idea from Edmonds and Moss (2004) to the modelling of human behaviour in ABSS where agent behaviour is often kept very simplistic.

In ABSS, computational frameworks are used to characterise agents and specify their decision-making process and behaviour. This research paper follows the classification and definitions of computational framework used in Li et al. (2008) which are described as follows. Computational frameworks can be divided into three overarching categories: Agent Frameworks, Multi-Agent Frameworks and Cognitive Architecture. Agent frameworks focus on individual agent behaviours from which social phenomena emerge and are used in agent design to determine decision-making and interaction processes of each agent. Multi-agent frameworks take a different perspective by specifying the relations between agent behaviour and collective behaviour in the modelling and simulation of multi-agent organisations. Cognitive Architectures are agent architectures based on cognitive science that help model social phenomena through the specification of agent cognition i.e. low-level behaviour in agents.

This research paper attempts to provide a guide for newcomers and researchers in the field of ABSS to the methods and challenges of modelling human behaviour in their social simulations. This will be accomplished by a literature review and an in-depth analysis and comparison of three representative computational frameworks. This will aid in structuring the different methods of modelling human behaviour and help to reach the potential of ABSS simulations. It differs from the work of Balke and Gilbert (2014) who performed a similar study in that the comparisons are made not only on the basis of agent characteristics but also relating to usability and implementation by researchers. The paper will explore these themes by answering the following research question:

What are the capabilities and limitations of modelling human behaviour using specific computational frameworks in ABSS research?

This article is structured as follows. Section 2 consists of the methodology used for the literature review, framework selection and formulation of comparison criteria. Section 3 summarises the chosen frameworks in the following order: the Belief-Desire-Intention (BDI) framework, the Normative Agent Architecture (NoA) and the Physical Conditions, Emotional State, Cognitive Capabilities, Social Status (PECS) framework. These frameworks are then evaluated and compared using specified criteria in Section 4. This is followed by a discussion in Section 5 on the results and the strengths and limitations of using these computational frameworks in ABSS research.

2 Methodology

2.1 Search Strategy

A first literature review was conducted on methods used for modelling human behaviour in ABSS. From this literature review, representative computational frameworks were selected for further study. This selection was based on the classification of computational frameworks used in Li et al. (2008). An agent framework, a multi-agent framework and a cognitive framework were chosen.

This research then focussed on studying three key frameworks for modelling human behaviour:

- the Belief-Desire-Intention (BDI) framework - an agent framework
- the Normative Agent Architecture (NoA) - a multi-agent framework
- the Physical Conditions, Emotional State, Cognitive Capabilities, Social Status (PECS) framework - a cognitive framework.

These frameworks were chosen for their significance in agent-based simulations and human behaviour modelling. They were researched and analysed for comparison using a set of relevant criteria adapted from the literature.

The search strategy was designed to capture relevant academic papers through a combination of database and journal-specific searches, supplemented by a snowballing technique. The database Scopus and search engine Google Scholar were the primary sources for identifying literature. Additionally, targeted searches were conducted on the websites of the Journal of Simulation and the Journal of Artificial Societies and Social Simulation to find domain-specific articles. A snowballing technique was also employed to find additional studies by reviewing the references of papers that were already identified as relevant. It was applied in particular to the work of Adam and Gaudou (2016) and Balke and Gilbert (2014) to determine other relevant literature. This method helped ensure a thorough exploration of existing literature on the BDI, PECS, and NoA frameworks.

The following keywords and search phrases were used in the Scopus database to identify relevant papers:

"bdi AND modelling AND human AND behaviour AND agent" (n= 96)
"belief AND desire AND intention AND framework AND agent AND review" (n= 11)
"beliefs AND desires AND intentions AND framework AND agent AND simulation" (n= 95)
"human AND behaviour AND pecs AND simulation AND modelling" (n= 8)
"human AND behaviour AND pecs AND simulation" (n= 11)
"human AND behaviour AND pecs AND modelling" (n= 12)
"noa AND behaviour AND simulation" (n= 8)
"human AND behaviour AND noa" (n= 22)
"noa AND multi-agent AND modelling" (n= 2)

These search terms were selected to specifically target papers that focus on agent-based modelling of human behaviour using the BDI, PECS, and NoA frameworks. The search process yielded a large number of studies. To ensure relevance, only papers written in English and containing the identified keywords either in the title or abstract were considered for review. Papers that did not meet these criteria were excluded from further analysis. The paper selection was a mix of theoretical studies and research that detailed how these frameworks were implemented in different scenarios and fields, so their applicability could be assessed in sections 4 and 5.

2.2 Model Comparison Criteria

Previous works in literature have sought to compare or evaluate different architectures for modelling human behaviour against a set of criteria. These criteria vary depending on the perspective and values of the researcher. Based on previous work in literature, a set of criteria were composed for use in the comparison of selected frameworks.

In Adam et al. (2017), the focus was on model comparison of two complex models for which different agent architectures had been utilised: Finite-State Machines(FSM) and BDI. Their criteria focussed on evaluating the agent architecture from the perspective of its complexity and usability to the researcher. Their selection criteria includes:

- Difficulty of description: Length of code (in characters)
- Difficulty of creation: Memory usage, Computational time
- Difficulty of appropriation: Understandability, Explainability, Extensibility
- Model credibility: error between model output and observed data

In Balke and Gilbert (2014), they explore models for agent decision-making through a discussion of 14 agent architectures. Its focus is on determining a suitable computational model of a human behaviour depending on the domain application. A list of 5 dimensions is used to differentiate agent architectures: Cognitive, Affective, Social, Norm Consideration and Learning. These dimensions differ from Adam et al. (2017) as they are based on desirable agent characteristics for ABSS problems and do not consider implementation aspects beyond the availability of resources and platforms.

Wray and Chong (2007) compares a number of computational tools for representing human behaviour. They differentiate between cognitive models which express low-level detail of human behaviour and human behaviour models which are often informed by psychological theories and used to generate more complex behaviour. Yet it is emphasised that cognitive models and human behaviour models share a continuous spectrum of characteristics when implemented.

A cognitive architecture is defined as an implementation platform for both cognitive models and human behaviour models. The work defines a set of general functionalities of cognitive architectures which are used to compare cognitive models and human behaviour models. These functionalities include:

- Context-based reasoning/action
- Least commitment/run-time decision-making
- Contextual conflict resolution
- Scalable knowledge bases
- Learning

Drawing from the literature on the comparison of agent architectures, existing criteria can be categorised at a higher level as focussing on agent capabilities, researcher usability and resources for implementation. Given that this work seeks to compare the architectures themselves and not the details of their implementations, not all of the aforementioned dimensions of comparison will be used.
Therefore, this work uses the following criteria in Table 1 to evaluate the selected agent architectures. The definitions of the criteria have been adapted from the papers discussed previously in this section.

Table 1: showing the comparison metrics for the study adapted from Balke and Gilbert (2014), Wray and Chong (2007), Adam et al. (2017)

	Criteria	Definition
Agent Capabilities	Context based reasoning	Agent's state trajectory is not prescribed but able to adapt based on perception of the current situation/context.
	Affective	Agents are able to have emotional representation by the architecture
	Learning	Agents are able to learn from past behaviour

	Norm consideration	Agents able to reason about social and legal norms, and consequently also can shape emergent norms
	Social	The architecture specifies social interactions or a means of agent communication
Implementation/Usability	Extensibility	Additional features can be introduced into the architecture with ease.
	Explainability of behaviour	The architecture can facilitate comprehension of the link between individual agent behaviours and outcomes.
	Understandability of architecture	Ease of interpretation and understanding of the architecture.
	Accessibility	Support and resources available to implement the computational architecture

The three frameworks were scored on all criteria to provide a comparative overview. The scoring was done based on the literature research on each of the frameworks. The following five words were used to provide the score for the criteria.

Table 2: Scoring terminology used to evaluate computational frameworks

Scoring Word	Definition
Good	Claimed by the developers and/or tested and proven in other research
Contested	Both claimed to be suited as well as unsuited by different researchers
Untested	Claimed to be appropriate, but untested in a research model
Limited	Claimed to be unsuited/difficult to do within the framework
Unspecified	No claims or tests performed

3 Frameworks

As mentioned in the methodology, the following frameworks were chosen for their significance in agent-based simulations and human behaviour modelling. BDI is a quite prominent theoretical frameworks in the field of simulating human behaviour whereas PECS was created to replace BDI, and NoA offers a slightly different perspective of a norm- governed society.

3.1 BDI: Beliefs Desires Intentions

Description

There exist several formalisms, architectures and implementations of the BDI model (Herzig et al., 2017; Norling, 2009; Rao & Georgeff, 1991; Wooldridge, 2000). As such, the following description of BDI is a generic overview based on the works by Adam and Gaudou (2016)and Norling (2009).

BDI architecture describes the behaviour of a rational agent in terms of mental attitudes of belief, desire and intention.

There are five main components of a BDI agent according to Norling (2009).

- ***Beliefs*** are the agent’s possibly flawed knowledge of the world i.e. their subjective reality.
- ***Desires*** (also referred to as ***goals***) consist of possible states of affairs that the agent seeks to achieve. To achieve these goals requires appropriate plans of action typically implemented as a plan library.
- A ***plan library*** is a set of actions, the corresponding desires they achieve and the conditions under which they operate.
- ***Intentions*** are the agent’s commitments to performing particular plans of action in order to achieve their desires.
- ***Reasoner*** is the engine that gathers information from the environment, updates the beliefs and desires, selects the plan to achieve the desire and selects the next action to execute from the set of intentions.

Adam and Gaudou (2016) summarise the core functionalities of a BDI architecture as requiring representations of beliefs, desires and intentions; rational and logical processes for selecting intentions and adaptable commitment to the set of current intentions.

At each time step, a BDI agent perceives its environment, updates its beliefs based on its perception, and pushes intentions generated by the reasoner into the relevant stack. For the next intention in the stack, the plan library is searched for all plans with preconditions agreeing with the agent’s beliefs and postconditions corresponding to the selected intention. These plans form a set of possible actions from which the agent chooses the most relevant actions depending on its state (Ye & Wang, 2017).

Application

The BDI model originated from philosophy with Michael Bratman's theory of intention. This theory was embraced by Artificial Intelligence (AI) researchers for the design of autonomous agents. From the AI field, BDI formalisms, software implementations and agent languages were created based on the theory(Herzig et al., 2017).

Though BDI is the most popular agent architecture for agent-based systems, in the field of Agent-based modelling and simulation, BDI remains underutilized (Adam & Gaudou, 2016). Yet BDI-based architectures have found applications for modelling decision-making processes in different fields such as the air traffic simulations (Wolfe et al., 2008), evacuation modelling (YenChern et al., 2021) and emergency response (Larsen, 2019).

Larsen (2019) stated that common agent-based simulation(ABS) platforms typically do not provide the ability to use agent architectures like BDI. However, of those that do provide the BDI framework, they incorporate BDI into the agent-based simulation in different ways. For instance, Repast, a popular ABS toolkit, embeds BDI systems into their ABM simulations by integrating the platform with the agent programming platform JACK(Padgham et al., 2011). Other platforms have extended their functionalities to enable BDI models like GAMA with its simple BDI architecture.

Challenges

BDI architecture does not account for many generic facets of human behaviour. As a result, when implemented, modellers who require additional features must take the time to do the implementation themselves (Norling, 2004).

Traditional BDI is known to lack specifications for: social interactions such as agent communication, normative considerations, emotions, a learning mechanism from past behaviour among other features(Balke & Gilbert, 2014). These deficiencies have resulted in the production of a range of frameworks that extend the basic BDI architecture to meet the needs of modellers such as eBDI and BOIDS architectures.

Norling (2004) points out that the assumed decision-making strategy of BDI agents is utility-based according to the original theory and contrasts with realistic decision-making behaviour as researchers have reported that an individual's decision-making strategy changes depending on their situation. A solution to this problem was formulated in Norling (2004) through an extended BDI architecture.

According to Adam and Gaudou (2016), BDI agents possess Theory of Mind i.e. the ability to reason about the thoughts and feelings (mental states) of themselves and other agents(Franchin, 2022). This contrasts with the work of Balke and Gilbert (2014) that maintains that traditional BDI does not implement this capability.

BDI does not perform well when implemented for an agent whose decision-making process is not well-defined or for which there is a lack of data (Wolfe et al., 2008). BDI requires granular data on individual behaviours which is often not readily available and has to be collected by other means such as participatory modelling or based on literature of psychological theories (Adam & Gaudou, 2016).

The BDI architecture lacks scalability due to the computational cost of implementation(Adam & Gaudou, 2016). Implementing BDI agents in agent-based simulations that often simulate large numbers of agents produces computational inefficiencies and severely slows simulation execution (Wolfe et al., 2008).

Conceptualising BDI agents can be difficult for non-programmers as they have to think about their model in terms of mental attitudes instead of objects and methods(Adam & Gaudou, 2016).

There exists a gap between the original BDI model and several of its implemented architectures in the role played by intentions. Intentions are high-level plans that are refined iteratively to result in an action. This process of intention refinement is described as fundamental to the BDI model in Herzig et al. (2017) which notes its absence from many studies on BDI agents.

Many BDI implementations (i.e. agent languages) are quite limited by their formalisation as they cannot be used to describe agents as having higher order beliefs (for instance, beliefs about what other agents believe) due to the lack of the formal logic semantics required for this task. Such BDI agents are unable to express social intelligence. (Herzig et al., 2017)

3.2 NoA: A Normative Agent Architecture

Description

NoA (Normative Agent Architecture) is an architectural framework designed to support the development of agent societies where the behaviours of agents are guided by norms, such as obligations, permissions, and prohibitions. These norms serve as directives for agents in contexts like automated business transactions within electronic commerce, ensuring correct contract execution and fostering trust between parties (Kollingbaum & Norman, 2004).

NoA operates through two main components: the NoA language and the NoA interpreter. The NoA language enables the specification of plans and norms, while agents within this architecture act based on a set of beliefs, pre-specified plans, and norms. Plans are executable when their preconditions are met, and norms are activated based on the agent's beliefs. Norms, which can define obligatory, permitted, or forbidden actions, are filtered through the NoA interpreter to ensure compliance (Kollingbaum & Norman, 2004).

Agents prioritise their own goals but must evaluate the impact of norms on their objectives before complying with them. If a norm conflicts with an important goal, an agent may choose to disregard it, though this typically results in penalties (López & Márquez, 2004). This flexibility is crucial for resolving conflicts between agents with differing goals within a synthetic society. Norms help mitigate these conflicts by ensuring that the behaviour of agents aligns with the overarching goals of the system (López & Márquez, 2004; Kollingbaum & Norman, 2003).

Norm adoption in NoA involves agents recognizing their responsibilities toward other agents and internalising norms that define these responsibilities. This process is guided by the agent's evaluation of its obligations, prohibitions, and permissions. Conflicts may arise between norms, particularly when obligations contradict prohibitions, but NoA provides mechanisms for resolving these issues through norm filtering and evaluation (Boissier et al., 2012; Kollingbaum & Norman, 2003). Obligations, Prohibitions and Permissions are defined as the following:

Obligations: Adopted under specific circumstances, these can conflict with other norms (Kollingbaum & Norman, 2003).
Prohibitions: These impose restrictions on the agent's behaviour.
Permissions: These override prohibitions, allowing the agent to engage in previously restricted actions (López & Márquez, 2004; Kollingbaum & Norman, 2004).

Further, a unique aspect of NoA is its treatment of agent autonomy. Autonomy in NoA refers to the agent's ability to choose between various plans to achieve its goals while still adhering to norms. This flexibility allows agents to adapt their behaviour in dynamic environments (López & Márquez, 2004). Agents in NoA can modify their plans based on changing conditions, which is particularly important in contexts like electronic commerce, where agents must respond to evolving norms and contractual obligations (Kollingbaum & Norman, 2003; Boissier et al., 2012).

Applications

NoA excels in managing complex decision-making processes where agents need to adhere to specific norms, such as in process-flow interactions and such. Its core strength lies in the way it handles obligations, prohibitions, and permissions, allowing agents to behave in predictable ways according to established rules. In human contexts, this could parallel behaviours motivated by laws, contracts, or societal expectations (López & Márquez, 2004). The architecture enables agents to reason about the consequences of their actions, allowing them to choose plans that not only satisfy immediate goals but also comply with the broader normative framework (Kollingbaum & Norman, 2003; López & Márquez, 2004).

The framework provides a robust framework for modelling human-like behaviour in multi-agent systems by incorporating norms, obligations, permissions, and prohibitions that guide agent actions. However, while NoA offers several advantages in simulating structured, rule-based behaviours found in human societies, it also has notable limitations in capturing the full complexity of human behaviour, as does any modelling technique.

NoA enables agents to dynamically evaluate conflicts between norms and personal goals. Agents can choose whether to violate a norm if it negatively affects a priority goal, much like how humans might prioritise personal objectives over social obligations (López & Márquez, 2004).

The implementation of NoA in literature, has demonstrated that agents can successfully navigate complex environments by adhering to predefined norms. The architecture supports adaptive behaviour, enabling agents to adjust their actions based on evolving norms, a critical feature for maintaining trust in agent interactions, as long as there are predefined norms for them to follow.

Challenges

While NoA is effective in managing norm-driven behaviours, it faces limitations when dealing with scenarios beyond pre-defined norms and plans. It struggles in dynamic, unpredictable environments where agents are required to operate outside established norms (Boissier et al., 2012; Kollingbaum & Norman, 2004). The following discussion highlights the key challenges that are a fundamental part of the NoA framework.

While NoA provides a strong foundation for modelling rule-based decision-making, it falls short in capturing the emotional aspects of human behaviour. Human decisions are often driven not only by rational calculations of costs and benefits but also by emotions such as guilt, shame, or pride, which are deeply intertwined with social norms (Argente et al., 2022). The absence of emotional modelling in NoA means it cannot fully simulate how humans might be motivated or constrained by feelings in response to norm violations or compliance. For instance, a human might obey a rule out of fear of social disapproval, a dynamic that NoA does not capture well (Argente et al., 2022). Following from this, NoA's reliance on predefined norms and plans also limits its ability to handle the unpredictability and emergent behaviours seen in human societies. Human behaviour is often shaped by spontaneous actions, cultural influences, or the development of new norms through social interaction, which NoA does not adequately simulate (Boissier et al., 2012).

Although NoA allows for some flexibility in plan selection and norm compliance, it ultimately operates within a structured, predefined set of actions and outcomes (Kollingbaum & Norman, 2004). It models agents as primarily self-interested entities with individual goals, which works well in structured systems like electronic commerce. However, human behaviour is often driven by complex social relationships, such as friendship, familial obligations, and long-term reciprocity, which influence decision-making beyond immediate goal satisfaction (Tufis & Ganascia, 2014). While NoA incorporates some aspects of social behaviour through norm compliance, it does not capture the subtleties of human relational dynamics, where norms may be negotiated or reinterpreted based on social ties.

3.3 PECS: Physical conditions, Emotional state, Cognitive Capabilities, Social Status

Description

PECS is a reference model introduced in 2000 by Urban that is supposed to support modellers in the design process of agent-based models that focus on human individual human decision making and human interactions (Urban, 2000). The model architecture enables agents to show reactive behaviour, deliberative behaviour as well as reflective behaviour (Schmidt, 2000a). This behaviour can be needed to create more realistic human behaviour of agents. The framework is especially strong when it comes to showing emergent behaviour resulting from a social environment and is therefore great for applications in a ABSS (Schmidt, 2000b.).

The framework architecture consists of a set of sub models that interact in specific ways. There are three layers. Firstly the input layer. The agents have sensors, which are connected to their perception. This layer is responsible for the processing of the input data. The second layer is composed of the set of internal state components. These components, Physis, Emotion, Cognition and Social Status are all connected in various ways and model the internal state of the agent. Lastly, these are connected to the last layer, the behaviour layer. This layer is responsible for the selection and execution of the agent's various actions (Urban, 2000)

With this architecture, modellers can be enabled to build in a wide range of human characteristics and behaviour. The various components can be made simple, or more complicated, depending on the modeller's needs and interests. Notable are the ability to build in reflective behaviour of agents. Based on the social status, emotional or physical condition, agents can be made to adapt their actions or become biassed. They can save information or have decaying memories due to age or other factors (Schimdt, 2000b)
All in all, there are many possibilities which lead to a wide range of possible applications in many simulation fields, where reactive, deliberative or reflective behaviour is required and where emergent behaviour is of interest. These fields include but are not limited to; evacuation simulation, social (-psychological) simulations or any simulation that requires human intelligent behaviour.

Application

Not a lot of applications of PECS were found in our literature review. Early implementations were made by the original research group to show the proof of concept. These include a simple survival model that includes emotion, food management and planning and a memory (Schmidt, 2000b). A second implementation they showed is to illustrate the possibilities to model group dynamics. Here students have to try to learn the most knowledge, by balancing their personal intelligence, as well as their social abilities to be able to join and function in a study group. They claim that the procedure they used 'show in exemplary and prototypical fashion the procedure to be followed in the modelling of complex human behaviour' (Schmidt, 2000b).

Implementations past the original proposal and proof of concept models are scarce and difficult to find. Furthermore, no built- in support for PECS in any popular ABSS simulation programs has been implemented. However, a few implementations have been made, like a more recent one from 2019 by Nguyen et al. They created an adapted framework from the PECS architecture inside a model of inhabitation. They mainly focused on the emotion component and used it as a part of their larger model. The emotional component is related to the behaviour component of the agents which controls how it makes decisions (Nguyen, 2019).

Challenges

The major challenge of PECS is the little research done after the initial proposal. Lots of potential has been shown, but up to now, based on the papers reviewed for this research, it has not been reached.

A second raised challenge comes from Nguyen et al. (2019). They call for more research to be able to expand their framework. They require more psychological and sociological research to be able to connect psychological motives to social groups' characteristics and to have more accurate input data. This could enable them to introduce more connected and accurate methods to their agents.

4 Comparison

In table 3, the three frameworks, BDI, NOA and PECS have been scored based on the criteria outlined in the methods. There are clear differences between the three frameworks.

Table 3: Comparison of three computational architectures: Belief-Desire-Intention (BDI), Normative Agent Architecture (NoA), Physis, Emotion, Cognition and Social Status (PECS) based on criteria specified in Table 1

	Criteria	BDI	NoA	PECS
Agent Capabilities	Context based reasoning	Good	Good	Good
	Affective	Unspecified	Untested	Good
	Learning	Unspecified	Unspecified	Good
	Norm Consideration	Unspecified	Good	Untested

	Social	Contested	Good	Good
Implementa-tion/ Usability	Extensibility	Good	Good	Good
	Explainability	Good	Good	Good
	Understanda-bility of Ar-chitecture	Good	Good	Good
	Accessibility	Contested	Limited	Limited

BDI scores the worst on the agent capabilities. This means that very specific behaviour and interactions that might be required have to be programmed by the researcher. However, in contrast to the other two frameworks, it scores contested on accessibility, so it could potentially be easier to implement using some available resources. Together this means that BDI could be a good framework to use for higher level implementations. NoA scores better than BDI on the agent capabilities. Norm consideration stands out, as it is the only framework that scores good for that criterion. Therefore, NoA could be a good choice when your agents and model have large interactions with norms. It is important to note that NoA scores limited on accessibility. That means it will be more complicated to implement the framework. PECS scores the best on the agent capabilities. It offers the widest options of the three frameworks for complex agents. Furthermore, just as NoA, it scores limited on accessibility. Therefore, the PECS frameworks should be considered when a complex agent is required, but similar to NoA, it will be relatively complicated to implement the framework.

5 Discussion

5.1 Table Evaluation

The basic BDI architecture was evaluated in Table 3 using the comparison criteria. Many aspects of human behaviour were seen to be unspecified in this architecture. This is compensated by the ease with which BDI can be extended to include additional features resulting in the development of a family of BDI-based architectures. On the other hand, BDI scored well for usability criteria as the architecture is based on folk psychology and expressed at a high-level of abstraction similar to the manner in which people explain their own actions making it easily comprehensible to researchers(Norling, 2009). Furthermore, BDI due to being a popular framework in the agent-based computing and artificial intelligence fields is supported by an extensive body of research and implementations. However, it is noted that though support for BDI is good in general, in ABSS, support for BDI is limited to a few platforms, therefore its accessibility was scored contested. As a result it is highly accessible for ABSS researchers who wish to learn about BDI but BDI implementation in ABSS tools may be limited.

NoA is great for modelling human behaviour that can be encapsulated though norms and interactions that can be guided by goals. As per the literature reviewed for this study, there are no experiments which describe the ability of the architecture to model or simulate emotions, but this has been stated as a theoretical possibility. As of the current research status, it fails to capture the complexity of human emotions- which are a critical part of human decision making. It is also not equipped to deal with unexpected situations which cause changes in the prescribed norms and goals of individual agents. It is unclear whether the goals of the individual agents can actually change contextually without prior-definition. However, it must be highlighted that this is an easy critique, and a critique that is true for most frameworks for modelling human behaviour, therefore it is important to reiterate at this point that the idea and purpose of these frameworks is not to replicate human behaviour, because that is an insurmountable task. Instead the goal is to identify the strengths and weaknesses of various frameworks, in an effort to aid modellers and scholars in deciding the most appropriate framework to use depending on the task at hand.

The evaluation of PECS in Table 3 showed a lot of strengths of PECS. It scored very well on most criteria. PECS does offer many strengths, especially when it comes to making a complex agent. However, as shown in the table, it has limited accessibility.

This comes from the little research and applications of PECS. This makes it very difficult to build an agent. There is little guidance and examples. Furthermore, because there is so much customizability, a lot of theories from psychology and sociology and very specific data are needed to make an accurate agent. Although there is significant research into the inner workings of humans, it still remains very difficult to find the right data (Wijermans, 2023).

5.2 Overlapping Limitations

From the research, certain challenges were observed to be common amongst the frameworks despite their differences. PECS, NoA and BDI require sufficient data on human behaviour to effectively use the architectures for agent design. This problem was also echoed in Kennedy (2012) as one the main challenges in modelling human behaviour in ABM. Another issue is the representation of affectivity which is difficult to encode and express in these frameworks as evidenced by both BDI and NoA. Although PECS includes the capacity for affectivity, this remains difficult due to limited accessibility and the need for the researcher to find and implement psychological theories of emotions. Limited implementation of these architectures in agent-based modelling platforms is another issue particularly for NoA and PECS although they differ from BDI which is the most popular architecture. As noted in (Wolfe et al., 2008), scalability of these architectures for use in ABMs can be computationally expensive. As such, there exists a trade-off between using these architectures to realise more accurate human behaviour modelling and the computational power available to researchers. The aforementioned challenges contribute to the ongoing underutilisation of these architectures in ABSS research to model more complex human behaviour.

6 Conclusion

This paper has provided insight into the capabilities and limitations of three specific computational frameworks that are used in ABSS research.

Common limitations were found to be availability of individual behavioural data, representation of emotions in agent decision-making, lack of support for architectures in ABSS platforms and computational cost.

The research has shown that for implementations of high-level behaviour, the BDI framework can be a good choice. It has relatively good support from applied research and a significant amount of literature to use. When it comes to modelling a system where norms and agent interactions are more important, it has been shown that NoA is a good pick. Normative architectures like NoA allow not only specification of the agent's micro-level behaviour but also incorporate top-down rules into the decision-making process by considering a potential action's compliance with norms, effect on other agents and effect on agent achieving its own goal. Lastly, PECS has been found to be able to offer the most customizability. It gives the modeller a lot of options and freedom to build their agent and make it complex. However, it also requires a lot of social and psychological theories and data, which can be a significant challenge.

It is important to note that this study was limited by the scope of the research and situated within the capabilities of the researchers. The scope was limited to the three frameworks of BDI, NoA and PECS but there are several computational frameworks that have been developed. Moreover, the paper's scope excluded testing and evaluating the few available implementations of these frameworks in software platforms. As such, this was not included in our criteria. It is recommended that future research could investigate the fidelity of architecture implementation against the conceptual models in platforms that claim to support them.

7 References

Adam, C., & Gaudou, B. (2016). BDI agents in social simulations: a survey. *The Knowledge Engineering Review, 31(3), 207-238.* https://doi.org/10.1017/S0269888916000096

Adam, C., Taillandier, P., & Dugdale, J. (2017). Comparing Agent Architectures in Social Simulation: *BDI Agents versus Finite-state Machines.* https://doi.org/10.24251/HICSS.2017.032

Argente, E., Del Val, E., Pérez-García, D., & Botti, V. (2022). Normative emotional agents: A viewpoint paper. *IEEE Transactions on Affective Computing, 13*(3), 1254-1267. https://doi.org/10.1109/TAFFC.2020.3028512

Balke, T., & Gilbert, N. (2014). How Do Agents Make Decisions? A Survey. *Journal of Artificial Societies and Social Simulation, 17(4), 13*. https://doi.org/10.18564/jasss.2687

Boissier, O., Bonnet, G., & Tessier, C. (2012). *Proceedings of the 1st Workshop on Rights and Duties of Autonomous Agents (RDA2) 2012. ECAI 2012.* Montpellier, France. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=9da72c3dd1845d8d3d1d05fc3c9123f8a3732c53

Davidsson, P. (2002). Agent Based Social Simulation: a Computer Science View. *Journal of Artificial Societies and Social Simulation, 5*. https://www.jasss.org/5/1/7.html

Edmonds, B., & Moss, S. (2004). From KISS to KIDS – *an 'anti-simplistic' modelling approach (Vol. 3415)*. https://doi.org/10.1007/978-3-540-32243-6_11

Franchin, L. (2022). Theory of Mind. In V. P. Glăveanu (Ed.), The Palgrave Encyclopedia of the Possible (pp. 1639-1644). *Springer International Publishing.* https://doi.org/10.1007/978-3-030-90913-0_3

Goldspink, C (2000). Modelling social systems as complex: Towards a social simulation meta-model. *Journal of Artificial Societies and Social Simulation vol. 3, no. 2.* https://www.jasss.org/3/2/1.html

Herzig, A., Lorini, E., Perrussel, L., & Xiao, Z. (2017). BDI Logics for BDI Architectures: Old Problems, New Perspectives. KI - Künstliche Intelligenz, 31(1), 73-83. https://doi.org/10.1007/s13218-016-0457-5

Kennedy, W. (2012). Modelling Human Behaviour in Agent-Based Models. In (pp. 167-179). https://doi.org/10.1007/978-90-481-8927-4_9

Kollingbaum, M. J., & Norman, T. J. (2003). Norm Consistency in Practical Reasoning Agents. *University of Aberdeen.* https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=10289dfda7356a0fac30f68c98120cc9c22f582d

Kollingbaum, M. J., & Norman, T. J. (2004). Strategies for Resolving Norm Conflict in Practical Reasoning. *University of Aberdeen.* https://www.researchgate.net/profile/Timothy-Norman-6/publication/247653878_Strategies_for_Resolving_Norm_Conict_in_Practical_Reasoning/links/53f5ece60cf2fceacc6f7ddb/Strategies-for-Resolving-Norm-Conict-in-Practical-Reasoning.pdf

Larsen, J. B. (2019). Going beyond BDI for agent-based simulation. *Journal of Information and Telecommunication, 3(4), 446-464.* https://doi.org/10.1080/24751839.2019.1620024

Li, X., Mao, W., Zeng, D., & Wang, F.-Y. (2008, 2008//). Agent-Based Social Simulation and Modeling in Social Computing. *Intelligence and Security Informatics, Berlin, Heidelberg*. https://doi.org/10.1007/978-3-540-69304-8_41

López, F. L. Y., & Márquez, A. A. (2004). An architecture for autonomous normative agents. *Proceedings of the Fifth Mexican International Conference in Computer Science. IEEE*. https://doi.org/10.1109/ENC.2004.1342594

Nguyen, B. V. D., Wang, T., & Peng, C. (2019). Integration of agent-based modelling of social-spatial processes in architectural parametric design. *Architectural Science Review*, *63*(2), 119–134. https://doi.org/10.1080/00038628.2019.1640107

Norling, E. (2004, 23-23 July 2004). Folk psychology for human modelling: extending the BDI paradigm. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, 2004. AAMAS 2004.* https://www.researchgate.net/profile/Emma-Norling/publication/2944539_Folk_Psychology_for_Human_Modelling_Extending_the_BDI_Paradigm/links/09e41508a4b1706b1d000000/Folk-Psychology-for-Human-Modelling-Extending-the-BDI-Paradigm.pdf

Norling, E. (2009). Modelling Human Behaviour with BDI Agents *[PhD thesis, The University of Melbourne].* https://rest.neptune-prod.its.unimelb.edu.au/server/api/core/bitstreams/14e50a72-5f2c-531e-bbf3-a536a7f2ec9f/content

Padgham, L., Scerri, D., Jayatilleke, G., & Hickmott, S. (2011, 11-14 Dec. 2011). Integrating BDI reasoning into agent based modeling and simulation. *Proceedings of the 2011 Winter Simulation Conference (WSC)*. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6147762

Rao, A. S., & Georgeff, M. P. (1991). Modeling rational agents within a BDI-architecture. *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning.*

Schmidt, B., Schneider, B. (2000a). AGENT- BASED MODELLING OF HUMAN ACTING, DECIDING AND BEHAVIOUR - THE REFERENCE MODEL PECS. *University of Passau.* https://scs-europe.net/services/esm2004/pdf/esm-55.pdf

Schmidt, B.(2000b). *The Modelling of Human Behaviour: Artificial Intelligence, Artificial Life, Psychology, Social Sciences*. https://jasss.soc.surrey.ac.uk/4/4/reviews/schmidt.html

Tufis, M., & Ganascia, J.-G. (2014). A Normative Extension for the BDI Agent Model. *In Proceedings of the 17th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines.* https://hal.sorbonne-universite.fr/hal-01495519v1/file/TufisGanascia_NormBDI_UKRE.pdf

Urban, C. (2000); PECS – A reference Model for the Simulation of Multi-Agent Systems, in: Suleiman, R., Troitzsch, K. G., Gilbert, G. N. https://doi.org/10.1007/978-3-642-51744-0_6

Wolfe, S. R., Sierhuis, M., & Jarvis, P. A. (2008). To BDI, or not to BDI: design choices in an agent-based traffic flow management simulation. *Proceedings of the 2008 Spring simulation multiconference, Ottawa, Canada.* https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=09e8161dce19af96d41c007d4fcf0c9345ebdb0e

Wooldridge, M. (2000). Reasoning about Rational Agents. *The MIT Press.* https://doi.org/10.7551/mitpress/5804.001.0001

Wray, R. E., & Chong, R. S. (2007). Comparing Cognitive Models and Human Behavior Models: Two Computational Tools for Expressing Human Behavior. *Journal of Aerospace Computing, Information, and Communication, 4(5), 836-852.* https://doi.org/10.2514/1.27099

Wijermans, N., Scholz, G., Chappin, É., Heppenstall, A., Filatova, T., Polhill, J. G., Semeniuk, C., & Stöppler, F. (2023). Agent decision-making: The Elephant in the Room - Enabling the justification of decision models fit in social-ecological models. *Environmental Modelling & Software*, *170*, 105850. https://doi.org/10.1016/j.envsoft.2023.105850

Ye, P., & Wang, S. (2017). A General Cognitive Architecture for Agent-Based Modeling in Artificial Societies. 5, 176-185. https://doi.org/10.1109/TCSS.2017.2777602

YenChern, N., Waishiang, C., KengWai, S., Khairuddin, M. A., Jali, N., & Mit, E. (2021). Developing fire evacuation simulation through BDI-based modelling and simulation. *Journal of Physics: Conference Series*, *2107*, *012047*. https://doi.org/10.1088/1742-6596/2107/1/012047