Computable Law as Argumentation-based MAS

Roberta Calegari[a], Andrea Omicini[b] and Giovanni Sartor[a]

[a] *Alma AI – Alma Mater Research Institute for Human-Centered Artificial Intelligence , Alma Mater Studiorum—Università di Bologna, Italy*
[b] *Dipartimento di Informatica – Scienza e Ingegneria (DISI), Alma Mater Studiorum—Università di Bologna, Italy*

Abstract

In this paper we sketch a vision of computable law as argumentation-based MAS, i.e., human-centred intelligent systems densely populated by agents (software or human) capable of understanding, arguing, and reporting, via factual assertions and arguments, about what is happening and what they can make possibly happen. A multi-agent system based on argumentation, dialogue, and conversation is, in this vision, the basis for making the law computable: through argumentation, dialogue, and adherence to social judgment, the behaviour of the intelligent system can be reached, shaped, and controlled with respect to the law. In such a scenario, computable law – and related intelligent behaviour – is likely to become associated with the capability of arguing about state and situation, by reaching a consensus on what is happening around and what is needed, and by triggering and orchestrating proper decentralised semantic conversations to decide how to collectively act in order to reach a future desirable state. Interpretability and explainability become important features for that sort of systems, based on the integration of logic-based and sub-symbolic techniques. Within this novel setting, MAS methodologies and technologies become the starting point to achieve computable law, even if they need to be adapted and extended for dealing with new challenges.

Accordingly, in this paper we discuss how this novel vision can build upon some readily-available technologies, and the research challenges it poses. We analyse a number of approaches and technologies that should be involved in the engineering of systems and services, and become core expertise for distributed systems engineers. Among the others, these include knowledge representation, machine learning, and logic argumentation.

Keywords

computable law, multi-agent system, argumentation, logic, hybrid approaches

1. Introduction

The research field of computable law studies the engineering of the law – i.e., the design of appropriate formal models and the development of the corresponding technology – to allow norms, terms and conditions to be represented in a machine-understandable way [1]. The aim is to enable machine and software agents to process, and reason about, legal abstractions with a certain degree of accuracy, and to take autonomous decisions based on this. In the context of computable law, two key factors have to be considered. The first is that a single or unique way

WOA 2020: Workshop "From Objects to Agents", September 14–16, 2020, Bologna, Italy
roberta.calegari@unibo.it (R. Calegari); andrea.omicini@unibo.it (A. Omicini); giovanni.sartor@unibo.it (G. Sartor)
http://robertacalegari.apice.unibo.it (R. Calegari); http://andreaomicini.apice.unibo.it (A. Omicini); https://giovannisartor.net/ (G. Sartor)
0000-0003-3794-2942 (R. Calegari); 0000-0002-6655-3869 (A. Omicini); 0000-0003-2210-0398 (G. Sartor)

of modelling legal knowledge cannot be taken for granted—namely, there are multiple ways of identifying and circumscribing the "law" to be modelled, and multiple ways of representing legal contents into automatically processable information structures. Therefore, the computable law model should be able to deal with heterogeneous "legal ontologies".

The second factor is that nowadays application scenarios for computable law are strongly characterised by the same "symbolic vs sub-symbolic" dichotomy that also informs the most promising techniques in artificial intelligence today. Indeed, if, on the one hand, legal intelligence has historically been bound to logic and argumentation (i.e., symbolic approaches), on the other hand, the applications of machine learning algorithms (i.e., sub-symbolic approaches) are increasing today, especially in the field of data analysis and predictive justice. The use of the latter techniques raises ethical and fairness issues linked to the lack of transparency and the possibility of hidden biases. Therefore, models aimed at making the law computable must somehow consider that dichotomy and should try a reconciliation – possibly via a blended integration – so to take advantage of each approach: benefiting from the strengths of each method, and smoothing the corresponding limits.

Along this line, in this paper we sketch a vision of computable law as an argumentation-based multi-agent system. As widely recognised, agent architecture is today the reference for the design of intelligent systems [2, 3, 4], it allows dealing with heterogeneous entities and models, and fits perfectly with the pervasive and distributed scenarios that the computable law aims at addressing. The vision we propose relies on the multi-agent system (MAS) abstractions and architecture [5] expanding them both from the point of view of norms and argumentation (as in existing works on normative MAS, see for instance [6, 7]), and of the integration between symbolic and sub-symbolic techniques.

In this scenario, the very nature of the system actors – intended as agents, but also as surrounding environments – embodies the concepts of computable law implementing and coordinating the activities of distributed processes in order to achieve either individual or common goals. In fact, computable law is likely to become associated with the capability of debating about situations and about the current context, by reaching a consensus on what is happening around and what is needed, and by triggering and directing proper decentralised semantic conversations to decide how to collectively act in order to reach the future desirable state of the affairs. Within this novel setting, interpretability and explainability become a remarkable feature of the multi-agent system.

1.1. Contributions of the Paper

Based on the envisioned scenario, the contributions of this paper are the following:

- We detail the concept of computable law as argumentation-based MAS (i.e., conversational agents), also with the help of a running example, and show how they affect the engineering of intelligent systems, challenging traditional approaches to distributed computing and calling for novel argumentation approaches.
- We investigate a number of approaches and technologies that should be involved in the engineering of systems and services in that original scenario, and should become core expertise for distributed systems engineering. Among the others, those include

knowledge representation and ontologies, machine learning, argumentation models and technologies, human-computer interfaces.

2. Motivating Scenario: self-driving cars

To ground the discussion, we first examine a case study in the area of traffic management, considering the near future of self-driving cars. In that scenario cars are capable of communicating with each other and with the road infrastructure while cities and roads are suitably enriched with sensors and virtual traffic signs able to dynamically interact with cars to provide information and supervision.

The choice of the case study is driven by the fact that both self-driving cars and the whole intelligent transportation domain – including traffic management – are a natural fit for computable law since they need real-time control and feedback – with respect to the norms and the current state of affairs – and should both adapt to legislation as well as to possible contingencies (also involving ethical choices and legal reasoning).

Accordingly, self-driving cars need to *(i)* exhibit some degree of intelligence for taking autonomous decisions, *(ii)* converse with the context that surrounds them, *(iii)* have humans in the loop, and *(iv)* be deeply intertwined with the law setting characterising the environment and the society. Generally speaking, the main success factor to address the goal of the car (i.e., reach a destination) is the capability to converse and dialectically interact with the surrounding environment (other vehicles, infrastructure, humans, etc.), in order to make the best choice and therefore actuate the best action.

Part of this scenario is already a reality in many (smart) cities around the world. There, IT technologies are exploited to improve both the organisation of the transportation infrastructure and the performance of some specific parts. Adaptive traffic lights, dynamic speed and flow metering, urban monitoring stations, and similar tools and algorithms are routinely employed by district administrations as a means to observe and revise traffic situations in (almost) real-time [8, 9], as well as to assist in urban planning.

In the following we analyse the scenario from a computational perspective – i.e., discussing in details actors and actions that come up as well as system requirements – in order to provide the bases for its reification on the most appropriate engineering approaches and corresponding architecture.

2.1. Analysis & requirements

In the envisioned desiderata scenario, passengers simply express their desire (e.g., “bring me to the hospital") and the car starts acting in autonomy, travelling towards the destination and without passengers to worry about the specific actions and decisions to be undertaken. There, autonomous goal-oriented smart agents (either software, objects, humans, etc.) are pervasive since multiple actors come into play, such as other self-driving cars or, for instance, autonomous intersection managers that regulate the flow of cars based on specific goals imposed by the municipality (e.g., reduce circulation in a specific area). Moreover, the environment is one of the main actors of such a system–providing context information, such as rules valid in that specific area or situation, i.e., encapsulating context knowledge and intelligence.

Looking in-depth at the expect behaviour of self-driving cars in an urban environment, it turns out that autonomous cars need to undertake a complex decision-making process, in compliance with norms and social rules, in almost the whole journey, which seamlessly integrates actions (and perceptions) at different levels.

The first action is the trip planning which must be done by considering the *global knowledge*—for instance, local laws, regulations, policies, and average traffic conditions of the intersections and roads along the path toward the destination, influenced by factors such as the day of the week, the hour of the day, the weather, and the like.

Planning is then modulated by considering all contingencies arising during the journey – such as, for instance, a car crash forcing a change path, a protest causing delays in our preferred path, etc. As a consequence, the original plan has to be adapted to the *local knowledge* (and perceptions) cars gather while enacting it; once again, the decision is closely related to legal issues: for instance, what is the best action to perform in order to avoid fines? Or, again, if an accident cannot be avoided, how to choose the ethically-preferable option?

Moreover, there are strict rules on *safety* we would like self-driving cars to automatically enforce, such as slowing down when passing nearby a school, breaking and setting aside the vehicle as soon as the horn of an ambulance is heard, etc. In other words, regardless of what the global and local knowledge may suggest, we abide by a set of general *commonsense rules* orthogonally considered valid.

Generally speaking, compliance with the legal rules by the autonomous cars – including also global, national, state, and local laws, regulations, and policies – must be guaranteed, unless contingent priority situations occur (such as when a norm has to be violated in order to save the life of a passenger/pedestrian).

The analysis carried out so far highlights some key requirements that engineering approaches and techniques have to fulfil. First of all, the envisioned scenario seamlessly integrates perceptions (and actions) at two different scales—namely, the macro and the micro.

The *macro level* of the system includes global knowledge and generally-valid rules, like universal norms and legal conventions, possibly modulated by commonsense reasoning. Moreover, with respect to the surrounding environment, the *macro level* deals with a *mid / long term* horizon and focus on the issue of traffic flow management–including, for instance, traffic flow forecasting and urban planning possibly learned from historical data analysis. On the other hand, the *micro level*, deals with the *short term* horizon, and mostly focuses on intersection management, including a few highly intertwined sub-problems—e.g. collision avoidance, minimisation of average delay, and congestion resolution. The macro-level and the micro-level act synergistically, exploiting some sort of integration in order to achieve individual and social goals.

As the last step, a suitable *infrastructure* for V2I (vehicle-to-infrastructure) and V2V (vehicle-to-vehicle) communication should be considered, for instance through the deployment of Road-side Units (RSU). This infrastructure should make it possible to convey information from the vehicle to the infrastructure and vice-versa. For instance, it should provide information about the road on which the vehicle is travelling and its specific rules, environmental conditions around the vehicle, traffic in the vicinity of the vehicle, and construction in the vicinity of the vehicle.

All the abovementioned ingredients should be mixed consistently in order to achieve the goal of the cars and therefore of the user, respecting current rules and convention. Therefore, the system knowledge base needs to be built by taking into account both macro and micro scales, and agents have to be able to reason and argue over it in order to achieve their goals.

In the following, we show that computable law naturally is an essential ingredient in a distributed multi-party conversation, or dialogue, based on distributed intelligence. It cannot be easily tackled with traditional approaches to distributed computing: instead, different approaches and techniques needs to be put in place and to be fruitfully integrated in order to meet the aforementioned requirements.

3. Computable law as conversation and distributed (micro-)intelligence in MAS

The wide variety of actors and requirements in the above-described scenario recall the MAS model and architecture as intrinsically suitable for addressing issues and challenges lay ahead. Along this line, the vision of computable law described in this paper is based on the fundamental roles of MAS – individuals (agents), society, and environment [10]– according to the Agents and Artefacts (A&A) meta-model [11].

Then, our vision enhances these fundamentals roles taking into account two more key concepts and related abstraction: norms and e-institutions – i.e., proper roles and abstractions for the normative environment – and micro-intelligence—i.e., rational reasoning capabilities to enhance both environment and individuals.

In a nutshell, the system is composed of several agents, each with his own personal goal to achieve. Agents' interaction and dialogue, on normative aspects and surrounding situations, allow them to complete their goals. For that reason, rules and conventions must have an active role in the system (e-institution) in order to be questioned, examined and respected. The system's actors – individuals, environment, e-institution – are intelligent (micro-intelligence) in that they can reason on knowledge and context, argue and explain the rationale behind their decision. The overall behaviour of the system transposes the social behaviour, the debate with the society can act as a social judgment on individuals and therefore can calibrate their individual actions.

Different roles of e-institution and micro-intelligence and their interaction are discussed below.

3.1. e-Institution

In the vision sketched above, particular attention has been given to *normative* concepts since norms and laws are the basic bricks upon which to build the concept of computable law, since individual and collective behaviours are both affected by norms. In particular, we leverage on the well-known concepts of electronic-institution (e-institution) – as in normative MAS [6] – and deliberative agents [12], setting up e-institutions via suitable coordination artefacts exploited as normative abstractions.

Loosely speaking, e-institutions are computational realisations of traditional institutions: i.e., coordination artefacts providing an environment where agents can interact according to stated laws – norms or conventions – in such a way that interactions within the e-institution reproduce norm-based interactions in the actual world. With the term deliberative agents, we emphasise the agents' autonomy stressed both by e-institutions and normative systems. Indeed, in such systems individuals possess the property of normative autonomy—i.e., can decide to violate a norm to achieve their goals, or to change their goals so that they match the existing norms. E-instutions can provide for real-time detection of violations, and norm-enforcement can be envisaged via different enforcement technique. For instance, a simple local blocking rule [13] can be applied (specific actions may be disallowed), or the application of the laws can be prioritised according to the contingent situation (some actions may be discouraged), or norms can be modelled in a game theory framework, where sanctions and rewards are provided for agents [14], or where agents are expected to cooperate according to norms that maximise the society's utility.

From an agent's point of view, two major benefits stem from modeling environments as e-institutions. On the one hand, e-institutions establish conventions – on behaviour, language, and protocols – that inform agents about the law and possibly induce them to comply. In a sense, the environment is given structure, so that the agents have an easy comprehension of its working laws. On the other hand, norm enforcer agents endowed with capabilities for acquiring norms dynamically and enforcing them in uncertain environments can be envisioned and spread all over the system. Recalling our motivating scenario: drivers or traffic wardens know when they can talk, what consequences their acts will have, and what actions are possible at each moment in time. These restrictions contains the set of actions that agents have to consider at each moment in time by limiting the set of options that agents have to think about.

From the systems properties' point of view, e-institution promote a clear embodiment of the laws that govern the system, thus making it observable and more explainable in its autonomous actions—i.e., the explicit role of institution allow to enforce a set of norms whose violation can be perfectly observed. With respect to the case study, e-institution incarnates global system knowledge as global, national, state, and local laws, regulations, and policies.

3.2. Micro-intelligence

In order to ensure different levels of knowledge – and intelligence – as highlighted in Section 2, the MAS model is extended with the concept of *micro-intelligence.* Micro-intelligence is exploited to manage precisely the micro-level of the system intelligence – i.e., local intelligence, as the intelligence of objects, things, but also of the environment – to be integrated with the macro-level intelligence. We recall the micro-intelligence definition from [15, 16] as the *externalised rationality* of cognitive agents, complementing their own in the sense of Clark and Chambers' *active externalism* [17], and under the perspective of Hutchins' *distributed cognition* [18], as depicted in Figure 1. In short, micro-intelligence is *external* because it does not strictly belong to agents, as it is a process independently executed by another entity – namely, an artefact – to whom the agent is (possibly, temporarily) coupled—in the sense of the distributed cognition's "extended mind". Moreover, it is *rational* because it is supposed to convey a sound inference and argumentation process in order to provide an explanation on the decision process. Micro-intelligence complements agents' own cognitive processes because it augments the cognitive capabilities of agents, by embodying situated knowledge about the local environment along with the relative inference and argumentation processes. Agents may even be unaware of the

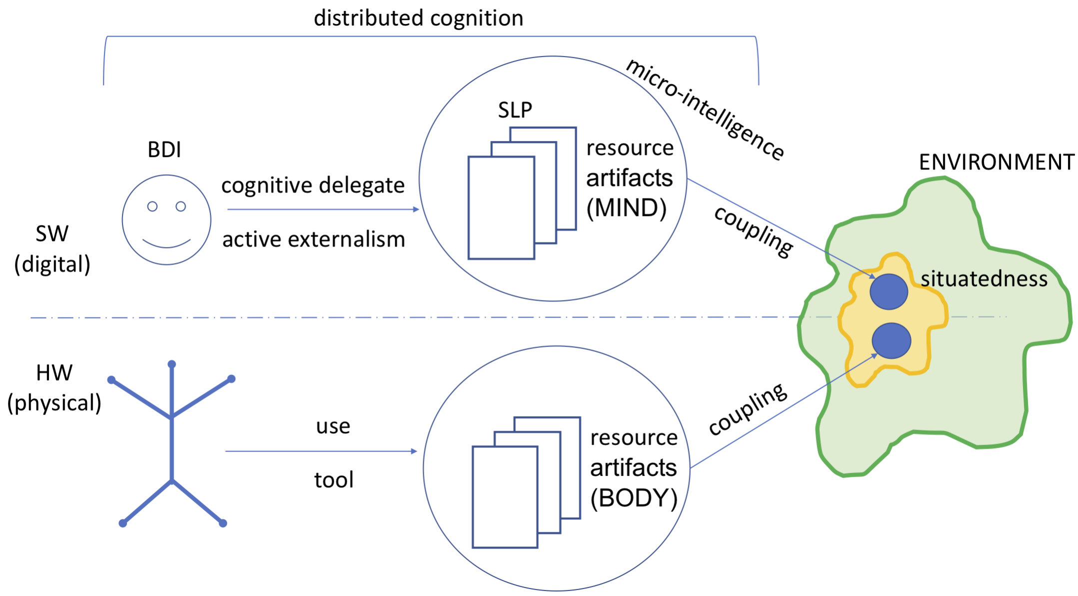

Figure 1: Micro-intelligence under the A&A perspective: distributed cognition is enabled by the active externalism of artefacts acting as cognitive delegates for the agent (extended) mind.

knowledge embodied in the environmental artefact delivering micro-intelligence. Under this perspective, artefacts act as delegates for intelligent (possibly, epistemic, but surely rational) behaviour, since they undertake inference processes on behalf of the interacting agents. Along this lines, our vision stems from two basic premises underpinning the above definition: *(i)* knowledge is locally scattered in a distributed environment, hence its situated nature; *(ii)* inference and argumentation capabilities are admissible and available over this knowledge, with the goal of extending the local knowledge through argumentation, induction, deduction, abduction, and the like.

Operationally, micro-intelligence is about scattering small chunks of machine intelligence all over a distributed and situated system, capable to enable the individual intelligence and the argumentation capability of any sort of devices. Micro-intelligence can be encapsulated in devices of any sort, making them smart, and capable to work together in groups, aggregates, societies. Thus, the micro-intelligence vision promotes ubiquitous distribution of intelligence in large pervasive systems such as those belonging to the pervasive and IoT landscape, in particular as a complement to agent-based technologies and methods, at both the individual and the collective level.

Note the micro-intelligence model – and related technologies – becomes fundamental also for the e-institution since it provides externalised rationality for reasoning and speaking with other system actors as well as the capability to reach properties such as interpretability and explainability, by leveraging on symbolic approach.

Moreover, micro-intelligence, as depicted in this summary, becomes the fundamental engine enabling conversation and argumentation between the system's actors and facilitating the control of contingent situations establishing the possibility to reach a consensus among speaking

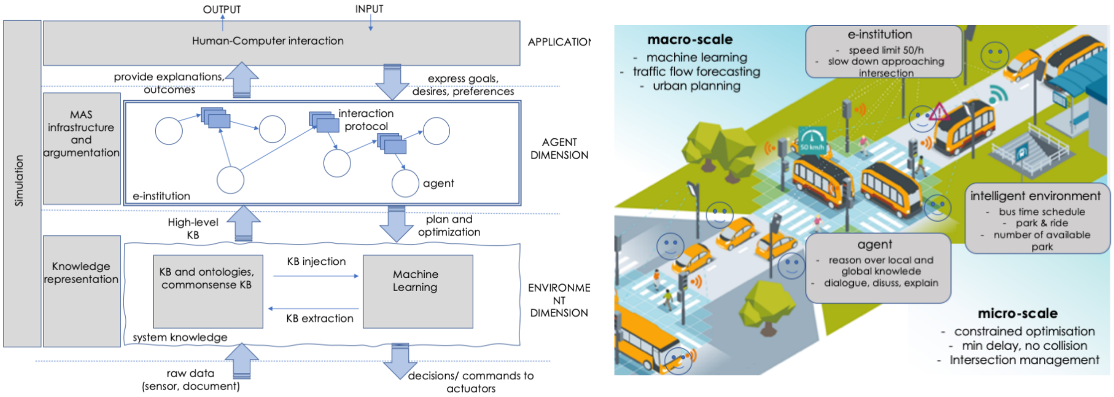

Figure 2: Main architecture components and techniques for realising the vision (left) and an exemplary deployment (right).

agents.

3.3. Overall vision

Mixing up all the above-mentioned ingredients, agents become capable of dialogue, exploiting a type of rational and symbolic intelligence that allows arguing and reasoning on the acquired knowledge. They become aware of the context – normative and social – in which they are immersed and through the argumentative process and the dialogue they can decide the next actions be implemented. The normative context is properly represented by the e-institution abstraction. Distributed (micro-)intelligence embodies the enabler of conversation between entities in the MAS infrastructure, making feasible the computation of the law and its incarnation in a computational system.

Dialogue is not expected only between agents but also with the environment and in particular with the e-institution. The bi-directional dialogue and interaction process, as well as enabling negotiation between system entities, also allow a prompt examination of each agent and of their status in order to detect automatically and real-time a possible deviation from the rules or policies envisaged. The issue of reaching a consensus in an ensemble of cooperating and interacting autonomous components by distributed negotiations has been already investigated in the field of multi-agent systems [19]. In particular, the argumentation-based negotiation area [20, 21] shows how argumentation can help in reaching global and individual goals and solutions, by letting agents converse and motivate their choices.

Figure 2 (left) summarises our vision by highlighting the main roles involved in the system as well as the two main activity flows—from data to users (left) and the opposite, from user goals and desires to activity planning, and lower-level commands (right). The grey boxes, that we describe more in detail in the following section, represent the technologies involved in the vision, while arrows represent the expected provided functionalities.

On one side, knowledge is collected from various sources – e.g., domain-specific knowledge, ontologies, sensors raw data – and is then exploited by agents that live in a normative environment (e-institutions). Agents' activity is then regulated by norms but also addicted to situated knowledge over which agents can reason and discuss in order to achieve goals. The multi-agent system, also thanks to its rational reasoning and argumentation capabilities, can provide outcomes to the users as well as explanations for their behaviours. On the other side, humans can insert input into the system – like desires, preferences, or goal to achieve – and these are transposed into agents' goal, corresponding activity planning, and lower-level commands for actuators.

The law, in this vision, become an internal component of the computational processes: legal norms, values, and principles are mapped and translated into, a computable representation of legal information – i.e., into the system knowledge – that is directly processed by computational entities. Agents become artificial legal agents able to comply with legal requirements and to reason over them.

Computable law assumes a vision of compliance/law by design in this architecture combining top-down compliance with predefined rules and bottom-up learning from cases and experience with capability to address regulatory conflicts according to legal values and principles.

4. Enabling technologies

The above considerations translate into different engineering approaches and therefore different nature of algorithms and techniques that can and must be considered when building the system. In the following, we encapsulate the main technologies involved as well as research challenges and opportunities.

4.1. Knowledge representation

Knowledge representation and related techniques are the cornerstone of the envisioned distributed system, to be able to argue and reason over it.

Of course, system knowledge has to take into account domain-specific knowledge and large-scale ontologies as repositories to interpret the knowledge bases available to the agents and to reason and argument over it. Knowledge could be continuously modified, adapted, and refined by the agents, according to their experience and perception of the environment or to learning from experience.

Accordingly, the knowledge base is plausible that is assembled by two main sources: on the one hand, ontologies and hand-crafted rules, on the other hand, rules learned from big data. Advances in machine learning will allow extracting knowledge from this data and to merge it with the former. Hybrid approach dealing with the integration of symbolic and sub-symbolic approaches becomes of paramount importance.

In this context, there are several issues and challenges to be tackled, to cite few, automatic extraction of knowledge from ML models, extraction of commonsense knowledge from the context, integration of the diverse knowledge in an appropriate logical language that allows argumentation and inference process to be performed. Several research fields are already facing these issues, but the general problem is far from being solved. For sure, we believe that a suitable integration of symbolic and sub-symbolic approaches can help in the achievement of the construction of proper system knowledge.

4.2. Machine learning

In our vision, a fundamental role is played by machine learning involved in different phases—namely, data processing & rule learning, and planning.

Data processing & rule learning. At the most straightforward level, machine learning techniques are clearly involved in raw input data elaboration, coming from sensors and/or documents, into more complex, high-level, structured information. Moreover, agents should be able to learn policies from past experience, by adapting both to the changing environment, and to the continuous progress of the society. Data aggregation, feature extraction, clustering, classification, data and pattern mining techniques are typically employed today to reach these objectives. We believe that hybrid approach could provide promising solutions to these tasks, by merging logic with probabilistic models and statistical learning, so to efficiently handle advantages of both symbolic and sub-symbolic approaches and moving towards explainable systems [22]. As highlighted above, the ML knowledge should somehow be translated into logical knowledge and properly merged with logical knowledge coming from ontologies or domain-expert norm translation or similar.

Planning. Distributed problem solving, planning, reinforcement learning, and cooperation [23] are some of the well-known ML techniques exploited in MAS. Our framework adds the challenge of integrating these techniques in the argumentation setting, so that the planning and cooperation derive from a continuous, natural interaction between agents with the environment. Once the user has specified his desires, the agent must be able to achieve them, interacting and coordinating with other individuals and with the e-institution to define the actions to perform and consequently defining appropriate plans to reify the decisions.

4.3. MAS & Normative MAS - middleware infrastructure

From a more implementation-oriented perspective, given that conversations are a new means of orchestrating the activities of distributed agents, an open research question – and a key one, too – is to understand which services should a middleware provide in order to support such distributed conversations.

The multi-agent infrastructure need not only to allow coordination among system actors but also include the possibility of customisable and reactive artefacts, capable of incorporating regulation and norms and micro-intelligence (possibly in the form of service). Moreover, the middleware should provide support for discussions via an open and shared discussion space, enabling dialogue among components that do not necessarily know each other in advance, and also providing services and or techniques for sharing knowledge, e.g., a tuple space [24]. However, unlike traditional tuple space models, the evolution of the conversation, the argumentation process and the reached consensus should be taken into account, also to be exploited in similar situations and/or to provide explanations. The best way to build such shared dialogue space, also taking into account different source of knowledge (e.g., commonsense kb, …) and different artefact acting both as law enforcer and intelligence promoter is a fertile ground for research.

4.4. Argumentation & logical reasoning

Argumentation is a necessary feature for agents to talk and discuss to reach an agreement. Several existing works establish the maturity of argumentation models as a key enabler of our vision [25, 26].

Despite the long history of research in argumentation and the many fundamental results achieved, much effort is still needed to effectively exploit argumentation in our envisioned framework. First, research on argumentation has mostly been theoretical, practical applications to real-world scenarios have only recently gained attention and are not yet reified in a ready-to-use technology [27]. Second, many open issues of existing argumentation frameworks regard their integration with contingency situation and situated reasoning to achieve a blended integration of the two concepts. Finally, the fundamental assumption every argumentation framework makes – that is, there must exist either an agreement among agents about the knowledge, or an external judge enacting some form of control over the argumentation process – is quite challenging preserving in our envisioned highly distributed, open, and dynamic scenario. The assumption is somehow related to the requirement of the formal model to have a coherent and logical conclusion, but neither of the assumptions is easy to have in a typical pervasive situation: reaching an agreement among many heterogeneous agents and devices is already a complex task, not obviously easily scalable, an external authority may be an unacceptable centralisation point. The argumentation architecture should be designed in order to be highly scalable, hybrid approaches should be investigated such as:

- provide many external authorities sharing the load of negotiating argumentations among a limited number of participants to enforce shared normative rules, possibly exploiting some notion of proximity;
- base the agreement on temporary agreement valid only for the duration of a conversation, defining somehow the concept of temporal locality;
- dually to the previous one, spatial locality may be the criterion to enforce partial consistency of normative and behavioural rules—according to the concept of "argumentation neighbourhood" where different distributed mediators act to enforce an agreement respecting the law.

In any case, we think that the concept of locality is crucial and should be considered along with distributed argumentation—coherently with the notion of micro-intelligence. In addition, it could be interesting to envision a framework where the specific argumentation and inference process can be swapped at runtime, making consideration on the specific situation, and introducing, for instance, abductive reasoning, or probabilistic argumentation or the algorithm deemed best and necessary for such a situation. The dialogue, therefore, becomes possible by following different rationales that can guide the decision-making process, possibly comparing different perspectives and visions.

4.5. Human-Computer interaction

Finally, techniques coming from natural language processing, computer vision speech recognition – already a reality in most everyday applications (e.g., Amazon Alexa, Google Home,…)

– become essential components of our vision for humans interaction. The challenge in the envisioned scenario is always related to the distributed issues, i.e., making commands possibly understandable to a multitude of agents and vice-versa. Existing algorithms should, therefore, be adapted for dealing with distributed and pervasive environments.

4.6. Simulation

Validation through simulation is of paramount importance when dealing with MAS, and even more when dealing on the integration of many different technologies—as in our vision. Simulation is concerned with the truthfulness of a model with respect to its problem domain and the correctness of its construction. In other words, simulation calls for verification, i.e., “building the system right" and validation, i.e., “building the right system" [28]. So, our vision leverages on consolidated simulation techniques that must become a core technology to validate and verify the final model and its properties.

In the traffic management scenario, for instance, by simulating car-following in continuous traffic flow and comparing simulation data with the data collected from the actual road, the reliability of the model and the architecture. Many simulation models have been proposed on the topic [29, 30, 31].

Also in this field, the simulation models must be able to mix ingredients from the techniques above discussed (symbolic and sub-symbolic) and should be adapted to meet the software engineering requirements these techniques refer to.

5. Conclusions

The heterogeneous nature of intelligence required by pervasive AI systems along with fears related to sub-symbolic techniques more and more exploited in such systems require for models and technologies guaranteeing explainability as one of the main requirement.

Within this context, we believe that argumentation can play a major role, as it allows debates to be studied and analysed, reasoning and persuasion to be exploited in dialogues, with a well-grounded theoretical framework.

In this paper, we show how different dialogues can occur in such pervasive contexts, highlighting the advantages of employing argumentation. Interpretability of decision making, tolerance to uncertainty, adaptiveness, robustness of the system, and improved trust by end-users and amongst interacting components, are the most notable benefits of the proposed approach.

The main limit of the argumentation approach is the typical assumption to have an external judge or authority that has to control the whole argumentation process, but this is very unlikely in a dynamic, distributed scenario like the one we propose. This aspect will certainly be the subject of future work.

However, model and techniques that should play a key role in the engineering of intelligent systems – even if with enhancements and extensions – have been discussed and constitute a starting point for further researches.

Acknowledgments

Roberta Calegari and Giovanni Sartor have been supported by the H2020 ERC Project "CompuLaw" (G.A. 833647). Andrea Omicini has been supported by the H2020 Project "AI4EU" (G.A. 825619).

References

[1] G. Sartor, P. Casanovas, N. Casellas, R. Rubino, Computable models of the law and ICT: State of the art and trends in european research, in: Computable Models of the Law, Springer, 2008, pp. 1–20. doi:10.1007/978-3-540-85569-9_1.

[2] M. J. Wooldridge, N. R. Jennings, Intelligent agents: Theory and practice, The Knowledge Engineering Review 10 (1995) 115–152. doi:10.1017/S0269888900008122.

[3] R. Calegari, G. Ciatto, E. Denti, A. Omicini, Logic-based technologies for intelligent systems: State of the art and perspectives, Information 11 (2020) 1–29. doi:10.3390/info11030167, Special Issue "10th Anniversary of Information—Emerging Research Challenges".

[4] F. Xhafa, S. Patnaik, M. Tavana, Advances in Intelligent Systems and Interactive Applications, volume 1084, Springer International Publishing, 2020. doi:10.1007/978-3-030-34387-3.

[5] C. Savaglio, M. Ganzha, M. Paprzycki, C. Bădică, M. Ivanović, G. Fortino, Agent-based internet of things: State-of-the-art and research challenges, Future Generation Computer Systems 102 (2020) 1038–1053. doi:10.1016/j.future.2019.09.016.

[6] G. Andrighetto, G. Governatori, P. Noriega, L. W. van der Torre, Normative multi-agent systems, volume 4 of *Dagstuhl Follow-Ups*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2013. URL: https://drops.dagstuhl.de/opus/volltexte/dfu-complete/dfu-vol4-complete.pdf.

[7] W. W. Vasconcelos, J. Sabater, C. Sierra, J. Querol, Skeleton-based agent development for electronic institutions, in: 1st International Joint Conference on Autonomous Agents and Multiagent Systems: Part 2 (AAMAS '02), ACM, New York, NY, USA, 2002, pp. 696–703. doi:10.1145/544862.544911.

[8] J. Rios-Torres, A. A. Malikopoulos, A survey on the coordination of connected and automated vehicles at intersections and merging at highway on-ramps, IEEE Transactions on Intelligent Transportation Systems 18 (2017) 1066–1077. doi:10.1109/TITS.2016.2600504.

[9] V. Gradinescu, C. Gorgorin, R. Diaconescu, V. Cristea, L. Iftode, Adaptive traffic lights using car-to-car communication, in: IEEE 65th Vehicular Technology Conference (VTC2007), 2007, pp. 21–25. doi:10.1109/VETECS.2007.17.

[10] A. Omicini, SODA: Societies and infrastructures in the analysis and design of agent-based systems, in: P. Ciancarini, M. J. Wooldridge (Eds.), Agent-Oriented Software Engineering, volume 1957 of *Lecture Notes in Computer Science*, Springer-Verlag, 2001, pp. 185–193. doi:10.1007/3-540-44564-1_12.

[11] A. Omicini, A. Ricci, M. Viroli, Artifacts in the A&A meta-model for multi-agent systems, Autonomous Agents and Multi-Agent Systems 17 (2008) 432–456. doi:10.1007/

s10458-008-9053-x, Special Issue on Foundations, Advanced Topics and Industrial Perspectives of Multi-Agent Systems.

[12] R. Rubino, A. Omicini, E. Denti, Computational institutions for modelling norm-regulated MAS: An approach based on coordination artifacts, in: O. Boissier, J. A. Padget, V. Dignum, G. Lindemann, E. Matson, S. Ossowski, J. S. Sichman, J. Vázquez-Salceda (Eds.), Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems, volume 3913 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2006, pp. 127–141. doi:10.1007/11775331_9.

[13] A. Perreau de Pinninck, C. Sierra, M. Schorlemmer, Distributed norm enforcement: Ostracism in open multi-agent systems, in: P. Casanovas, G. Sartor, N. Casellas, R. Rubino (Eds.), Computable Models of the Law, Springer, 2008, pp. 275–290. doi:10.1007/978-3-540-85569-9_18.

[14] L. Denant-Boemont, D. Masclet, C. N. Noussair, Punishment, counterpunishment and sanction enforcement in a social dilemma experiment, Economic Theory 33 (2007) 145–167. doi:10.1007/s00199-007-0212-0.

[15] A. Omicini, R. Calegari, Injecting (micro)intelligence in the IoT: Logic-based approaches for (M)MAS, in: D. Lin, T. Ishida, F. Zambonelli, I. Noda (Eds.), Massively Multi-Agent Systems II, volume 11422 of *Lecture Notes in Computer Science*, Springer, 2019, pp. 21–35. doi:10.1007/978-3-030-20937-7_2.

[16] R. Calegari, Micro-Intelligence for the IoT: Logic-based Models and Technologies, Ph.D. thesis, Alma Mater Studiorum—Università di Bologna, Bologna, Italy, 2018. doi:10.6092/unibo/amsdottorato/8521.

[17] A. Clark, D. J. Chalmers, The extended mind, Analysis 58 (1998) 7–19. URL: https://www.jstor.org/stable/3328150.

[18] J. Hollan, E. Hutchins, D. Kirsh, Distributed cognition: Toward a new foundation for human-computer interaction research, ACM Transactions on Computer-Human Interaction (TOCHI) 7 (2000) 174–196. doi:10.1145/353485.353487.

[19] S. Kraus, R. C. Arkin, Strategic Negotiation in Multiagent Environments, Intelligent Robotics and Autonomous Agents, MIT Press, 2001. URL: https://mitpress.mit.edu/books/strategic-negotiation-multiagent-environments.

[20] L. Amgoud, Y. Dimopoulos, P. Moraitis, A unified and general framework for argumentation-based negotiation, in: 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007), ACM, 2007, pp. 1–8. doi:10.1145/1329125.1329317.

[21] I. Rahwan, S. D. Ramchurn, N. R. Jennings, P. Mcburney, S. Parsons, L. Sonenberg, Argumentation-based negotiation, The Knowledge Engineering Review 18 (2003) 343–375. doi:10.1017/S0269888904000098.

[22] R. Calegari, G. Ciatto, A. Omicini, On the integration of symbolic and sub-symbolic techniques for XAI: A survey, Intelligenza Artificiale 14 (2020) 7–32. doi:10.3233/IA-190036.

[23] P. Stone, M. Veloso, Multiagent systems: A survey from a machine learning perspective, Autonomous Robots 8 (2000) 345–383. doi:10.1023/A:1008942012299.

[24] S. Mariani, A. Omicini, Coordination in situated systems: Engineering MAS environment in TuCSoN, in: G. Fortino, G. Di Fatta, W. Li, S. Ochoa, A. Cuzzocrea, M. Pathan (Eds.), Internet and Distributed Computing Systems, volume 8729 of *Lec-*

ture Notes in Computer Science, Springer International Publishing, 2014, pp. 99–110. doi:10.1007/978-3-319-11692-1_9.

[25] J. Hulstijn, L. W. van der Torre, Combining goal generation and planning in an argumentation framework, in: International Workshop on Non-monotonic Reasoning (NMR'04), 2004, pp. 212–218.

[26] S. Modgil, M. Caminada, Proof theories and algorithms for abstract argumentation frameworks, in: Argumentation in Artificial Intelligence, Springer, 2009, pp. 105–129. doi:10.1007/978-0-387-98197-0_6.

[27] R. Calegari, G. Contissa, F. Lagioia, A. Omicini, G. Sartor, Defeasible systems in legal reasoning: A comparative assessment, in: M. Araszkiewicz, V. Rodríguez-Doncel (Eds.), Legal Knowledge and Information Systems. JURIX 2019: The Thirty-second Annual Conference, volume 322 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2019, pp. 169–174. doi:10.3233/FAIA190320.

[28] D. Calvaresi, G. Albanese, J.-P. Calbimonte, M. Schumacher, SEAMLESS: Simulation and analysis for multi-agent system in time-constrained environments, in: Advances in Practical Applications of Agents, Multi-Agent Systems, and Trustworthiness. The PAAMS Collection, Springer International Publishing, Cham, 2020, pp. 392–397. doi:10.1007/978-3-030-49778-1_30.

[29] M. Mashayekhi, H. Du, G. F. List, M. P. Singh, Silk: A simulation study of regulating open normative multiagent systems., in: 25th International Joint Conference on Artificial Intelligence (IJCAI-16), 2016, pp. 373–379. URL: https://www.ijcai.org/Proceedings/16/Papers/060.pdf.

[30] J. Dai, X. Li, Multi-agent systems for simulating traffic behaviors, Chinese Science Bulletin 55 (2010) 293–300. doi:10.1007/s11434-009-0230-3.

[31] F. Zhang, J. Li, Q. Zhao, Single-lane traffic simulation with multi-agent system, in: 2005 IEEE Intelligent Transportation Systems, IEEE, 2005, pp. 56–60. doi:10.1109/ITSC.2005.1520219.