A Logical Analysis of Mamdani-type Fuzzy Inference, I Theoretical Bases

Simone Bova, Pietro Codara, Daniele Maccari and Vincenzo Marra

***Abstract*— This paper is divided into two parts. In the present Part I, our main objective is to analyse Mamdani-type fuzzy control systems in logical terms, with special emphasis on the fuzzy inference process. To that end, we provide our own inference procedure, cast in the language of standard many-valued logics. We give an ample discussion of the logical meaning of our procedure. We eventually show how to fully recover Mamdani-type fuzzy inference from the latter. In this sense, then, our proposal may be regarded as a logical interpretation of Mamdani-type fuzzy inference. In Part II of this paper, we report on the results of an experiment on the technical analysis of the financial markets based on fuzzy techniques. The core algorithm implements the inference procedure described in this first part of the paper. In Part II, we will argue that the experimental results support the claim that our present theoretical analysis provides a sound interpretation of Mamdani-type fuzzy inference.**

I. Introduction

In 1975, Mamdani and Assilian published their influential paper [1] on what are nowadays known as *Mamdani-type fuzzy control systems*. Their work was inspired by an equally influential publication by Zadeh [2]. Interest in fuzzy control has continued ever since, and the literature on the subject has grown rapidly. A survey of the field as of 1990 with fairly extensive references may be found in [3]. A more recent perspective with an eye to future challenges is [4].

Along with interest in the manifold issues related to the theory of control proper, closer attention to the more theoretical, genuinely logical aspects of fuzzy control began to emerge. While this is not the place to provide an account of the relevant literature, it is appropriate to recall the landmark work of Hájek [5]. With hindsight, it was Hájek's book that did the most to get a number of professional logicians involved with the logical aspects of fuzzy control. The present paper, divided into two parts – this one, and [6] – falls into this line of investigation.

In Hájek's approach, fuzzy logic is formalised by a many-valued logical system – at least when the term 'logic' is interpreted in a sufficiently narrow sense. Many-valued logic, in turn, has a long-standing tradition of its own. To mention just one example, the well-known Łukasiewicz logic arose out of the efforts of Jan Łukasiewicz in the Twenties and Thirties to provide a philosophical rebuttal of determinism and the related problem of future contingents. With time, it has grown into a well-developed chapter of mathematical logic, with thriving connection to several fields of mathematics and computer science; for a comprehensive treatment, including historical references, see the standard handbook [7].

A third largely independent line of development that may be mentioned here is the philosophy of vagueness. Here, philosophers have been developing for decades sophisticated analyses of the semantics of inherently vague statements, such as e.g. "The ball is red". Some of their theories countenance degrees of truth, and are in fact tightly related to many-valued – and therefore fuzzy – logic; others reject the use of such degrees for the analysis of vagueness. The interested reader will find [8] a useful starting point.

Returning to the present paper, our main objective in this first part is to provide a first attempt at our own logical analysis of fuzzy control systems in the style of Mamdani, with special emphasis on the fuzzy inference process. In [5, Ch. 7], after having developed a unified theory of several available many-valued logical systems, Hájek provides his account of what is usually termed approximate inference. Among other things, he deals with the fuzzy inference process that essentially goes back to Mamdani's original paper [1]. Mamdani's inference process has evolved into a standard tool of fuzzy control; it is available, for instance, in the widespread Fuzzy Logic Toolbox of MATLAB. We shall throughout call it *Mamdani-type* (*fuzzy*) *inference*. Although we cannot discuss Hájek's treatment in detail here, we recall that it is based on quantified, predicate logic. By contrast, what we present in this paper is only based on the more elementary notion of propositional logic. In other words, we will never need to formalise such statements as "For all x, x is a red ball", which involve the (fuzzy, or many-valued) predicate $\mathrm{Red}(x)$. Rather, it will be enough for us to deal with the simpler (fuzzy, or many-valued) proposition "The ball is red", where it will be understood that the proposition refers to a specific, given ball (which may or may not be red, or red to a degree.) It may be argued that this is closer in spirit to the way Mamdani-type inference features in actual systems. More generally, our aim in this first part of the paper is to provide a recasting of Mamdani-type inference in purely logical terms, using standard notions from the tradition of mathematical logic. As will become clear, we carry out our analysis mainly on the semantical side; further syntactical or proof-theoretic considerations are left to future research. In particular, this entails that we will not attempt to interpret Mamdani-type inference in terms of the syntactic

S. Bova is with the Department of Mathematics, Vanderbilt University, 1326 Stevenson Center, Nashville, TN 37240-001, USA (email: simone.bova@vanderbilt.edu).

P. Codara, D. Maccari, and V. Marra are with the Dipartimento di Informatica e Comunicazione, Università degli Studi di Milano, via Comelico 39, I-20135 Milano, Italy (email: codara@dico.unimi.it, daniele.maccari@gmail.com, marra@dico.unimi.it).

consequence relation of a given many-valued logic, although the work presented here can be read as a preliminary step in that direction.

Before discussing the organisation of the first part of this paper, we address a remark by one of the anonymous referees. It is well known that, in the literature, Mamdani-type systems are often regarded as mere interpolation tools. In fact, it is often argued that Mamdani-type inference has nothing to do with logical inference proper, so that the former is a misnomer. See e.g. [9], [10] and references therein. While our paper does not claim to reinterpret Mamdani-type 'inference' as genuine logical inference, it does show how to recast it in *purely logical terms*, as mentioned above. Hence, our results do something to redress the balance of the widespread view that regards Mamdani-type inference as an interpolation tool largely devoid of logical significance.

The contents of the first part of this paper are as follows. In Sect. II, we provide the needed background on propositional many-valued logics. In Sect. III we analyse the main components of a fuzzy control system from the point of view adopted in this paper. We deal in turn with fuzzification, fuzzy inference, and defuzzification. Our main focus is on fuzzy inference; our treatment of fuzzification and defuzzification, being instrumental to fuzzy inference, is necessarily curt. A similar but more emphatic remark applies to control theory proper: we do not even touch upon significant, well-known issues (such as stability) usually central to control theory. In Sect. III-B.1 we give a detailed account of Mamdani-type fuzzy inference. Parallel to that, in Sect. III-B.2 we describe our own alternative, logic-based procedure. We follow that up in Sect. IV with an extensive discussion of the logical significance of the procedure spelled out in Sect. III-B.2. Finally, in Sect. V, we show how to fully recover Mamdani-type inference from our own procedure, thus showing that the former is a special case of the latter. To all effects, then, our proposal in Sect. III-B.2 may be regarded as a logical interpretation of Mamdani-type fuzzy inference.

While we believe that theoretical considerations such as the ones put forth in this paper are important, we also believe that they should eventually be tested against experience. The second part of this paper [6] takes a step in this direction. In Part II, we report on the results of an experiment on the technical analysis of the financial markets based on fuzzy techniques. The core algorithm implements our procedure described in Sect. III-B.2 of this first part of the paper; further background and a discussion of the results are postponed to Part II.

II. Many-valued Logics Based on Triangular Norms

Quite generally, a propositional logic $\mathscr{L}$ is determined by a syntax, a semantics, and an inferential mechanism that ties the two together. It will be expedient here to recall the details in the case of classical (or Boolean) logic.

For the syntax, we start with a set V of *propositional variables*, or *atomic formulas*, that represent those propositions that cannot – or that we do not wish to – analyse into simpler constituents. To fix ideas, say

$$\mathrm{V} = \{X_1, X_2, \ldots, X_n, \ldots\} \ . \tag{*}$$

The set V is the *language* of $\mathscr{L}$. For our purposes here, a finite language will always suffice. To (*) we adjoin two symbols, say $\top$ (the *verum*) and $\perp$ (the *falsum*), that stand for a proposition that is always true and one that is always false, respectively; these are the *logical constants* in the language. To model compound propositions, we use the *logical connectives* $\wedge$ (for conjunction), $\vee$ (for disjunction), $\rightarrow$ (for implication), and $\neg$ (for negation). The usual inductive definition of a formula now reads as follows.

- All propositional variables are formulas, and so are $\top$ and $\perp$.
- If φ and ψ are formulas, then so are $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $(\neg\varphi)$.
- Nothing else is a formula.

(In practice, following well-known conventions on the precedence rules among connectives, redundant parentheses are usually omitted from formulas.) Let us write FORM for the set of formulas constructed over the language V.

Turning to semantics, in classical logic the meaning of a formula is assumed to coincide with its *truth value*. It is further assumed that each formula has two possible truth values only, say 0 – representing falsehood – and 1 – representing truth. Accordingly, we consider (*truth-value*) *assignments* to formulas, namely, functions $\mu\colon \mathrm{FORM} \rightarrow \{0, 1\}$ subject to the familiar conditions:

1) $\mu(\top) = 1$, $\mu(\perp) = 0$;
2) $\mu(\varphi \wedge \psi) = 1$ iff both $\mu(\varphi) = 1$ and $\mu(\psi) = 1$;
3) $\mu(\varphi \vee \psi) = 1$ iff either $\mu(\varphi) = 1$ or $\mu(\psi) = 1$, or both;
4) $\mu(\varphi \rightarrow \psi) = 0$ iff $\mu(\varphi) = 1$ and $\mu(\psi) = 0$;
5) $\mu(\neg\varphi) = 1$ iff $\mu(\varphi) = 0$,

for all $\varphi, \psi \in \mathrm{FORM}$. Granted that 0 and 1 stand for falsehood and truth, conditions 1)–5) fix a precise meaning for the logical connectives. Thus, for instance, 4) prescribes that $\rightarrow$ is to be interpreted as *material implication* – notoriously only a very rough approximation of its natural language counterpart "*If* ... *then* ...", but nonetheless useful. For future reference, let us note that in classical logic there exists a crucial relationship between conjunction and implication. Let μ be any assignment, and consider formulas φ, ψ, and χ. Then

$$\mu(\varphi \wedge \chi) \leq \mu(\psi) \text{ if and only if } \mu(\chi) \leq \mu(\varphi \rightarrow \psi) \ . \tag{**}$$

Relationship (**) is known as *residuation*. Speaking informally, it says that the truth value of the implication $\varphi \rightarrow \psi$ is as large as it can be, provided one asks that its conjunction with the truth value of the antecedent φ does not exceed the truth value of the consequent ψ. We shall return to (**) shortly.

A formula is a *tautology* if it takes value 1 (the *designated truth value*) under each assignment. We can conceive of an assignment μ as a *possible world*, determined by the fact that each sentence φ utterable in that world is known to be either true ($\mu(\varphi) = 1$) or false ($\mu(\varphi) = 0$). On this account, tautologies are just those formulas that are true in every possible world, i.e. necessarily true by virtue of their form alone, not their specific content. A well-known example is the *tertium non datur* principle: for any φ, $\varphi \vee \neg\varphi$ must be true under any circumstance (i.e. for every choice of μ), independently of the specific content of φ.

To link syntax and semantics, one selects an appropriate subset of formulas $\text{AX} \subseteq \text{FORM}_\text{V}$, called the *axioms* of the logic. In the classical case, the set AX can be effectively described via a finite list of *axiom schemata* such as e.g. $\varphi \to (\psi \to \varphi)$; AX is then the (infinite) set of all formulas obtainable from such schemata upon instantiating φ and ψ (in the previous example) with two specific formulas. (Observe that, on this definition, AX is a *decidable* subset of FORM_V: there is an algorithm that on input $\varphi \in \text{FORM}_\text{V}$ outputs YES if $\varphi \in \text{FORM}_\text{V}$, and NO otherwise.) Finally, a formula φ is *provable* (from AX) if there is a finite list $\psi_1, \psi_2, \ldots, \psi_n = \varphi \in \text{FORM}_\text{V}$ such that each ψ_i either lies in AX, or is obtainable from two previous formulas in the list via the *modus ponens* deduction rule: given formulas α and $\alpha \to \beta$, infer β. The *completeness theorem* shows that syntax and semantics match perfectly: *A formula is provable if and only if it is a tautology*. (The easy half of this equivalence – the fact that only tautologies are provable – is more accurately called *soundness*. We are going to call 'completeness theorem' the whole statement, for short.) The following generalisation of the completeness theorem is important. A *theory* is any set Θ of additional *extra-logical* axiom schemata that encode the assumptions required to hold in a specific application domain of interest. A formula φ is a *syntactic consequence* of Θ if it is provable from $\text{AX} \cup \Theta$ in the same sense as before; and it is a *semantic consequence* of Θ if each assignment that evaluates each instance of a schema in Θ to 1, also evaluates φ to 1. Now the *completeness theorem for theories* reads: *For every theory Θ, the syntactic and semantic consequences of Θ coincide*.

The time-honoured machinery above has of course proved its worth in several domains, both theoretical and applied. As things stand, its scope is restricted to crisp, yes/no sentences; can we extend it to encompass *fuzzy* sentences? Among a number of alternative frameworks that arguably provide affirmative answers to that question, we give an account of the highly influential one systematised by Hájek [5] in the late Nineties.

Suppose we agree on the following basic assumptions about $\mathscr{L}$.

1) $\mathscr{L}$ has precisely the same syntax (cf. V and FORM) as classical logic, except that we only take $\wedge$, $\to$, $\bot$, and $\top$ as primitive connectives, and we regard $\vee$ and $\neg$ as derived connectives whenever possible. While for classical logic this is an immaterial assumption – one can define $\neg\varphi \equiv \varphi \to \bot$ and $\varphi \vee \psi = \neg(\neg\varphi \wedge \neg\psi)$ – in a more general setting this is not always the case; more on this below.
2) Possible worlds are *fuzzy*, namely, assignments $\mu\colon \text{FORM}_\text{V} \to [0,1]$ take values in the real unit interval $[0,1]$.
3) Absolute truth and falsehood are retained; specifically, $\mu(\bot) = 0$ and $\mu(\top) = 1$ hold for any assignment μ.
4) As in the classical case, $\mathscr{L}$ is *truth-functional*. That is, for each assignment μ and each formula φ, the value $\mu(\varphi)$ only depends on the values that μ assigns to those atomic propositions in V that occur in φ. Equivalently, this means that the semantical interpretation of the connectives $\wedge$ and $\to$ is given by appropriate two-place functions $f_\wedge$, $f_\to\colon [0,1] \times [0,1] \to [0,1]$.
5) As a binary operation on $[0,1]$, $f_\wedge$ is associative, commutative, non-decreasing in both arguments, and continuous. These assumptions are supposed[1] to reflect common usage of conjunctions of vague sentences in natural language.
6) For all $x, y \in [0,1]$, $f_\wedge(x,1) = x$, and $f_\to(x,y) = 1$ if and only if $x \leq y$, so that the restrictions of $f_\wedge$ and $f_\to$ to $\{0,1\}^2$ coincide with the interpretation of classical conjunction and implication.
7) Conjunction and implication are tied together by the same residuation relationship (**) as in the classical case. With hindsight, it can be shown that this assumption guarantees that the fuzzy version of the *modus ponens* deduction rule is as powerful as it can be if it is to be sound. For a detailed discussion of this point, we refer to [5, 2.1.3].

In this scenario, a binary operation known as a *triangular norm* (*t-norm* for short), provides a suitable interpretation for fuzzy conjunction. Indeed, a t-norm[2] $*$ is a continuous binary function on $[0,1]$ that is associative, commutative, monotone ($x \leq y$ implies $x * z \leq y * z$) and has 1 as unit ($x * 1 = x$). Importantly, it turns out that *each t-norm has a unique associated binary operation $\to^*\colon [0,1] \times [0,1] \to [0,1]$ such that the condition*

$$x \to^* y = \max\{z \mid x * z \leq y\} \qquad (**')$$

is satisfied. The upshot of this is that if we then interpret $\wedge$ through $*$ and $\to$ through $\to^*$, then because of (**′) the residuation relationship (**) will continue to hold in the fuzzy setting, too, as required by our assumption 7). Notice that $x \leq y$ is equivalent to $x \to^* y = 1$, as required by 6).

Given a t-norm $*$, its associated *triangular conorm* (*t-conorm* for short), written $+^*$, is semantically defined via a generalised form of the De Morgan Laws, setting $x +^* y = 1 - ((1-x) * (1-y))$. While many researchers consider $f_\neg(x) = 1 - x$ to be the natural interpretation for fuzzy negation, a syntactic counterpart of $f_\neg$ is not always available within the logic $\mathscr{L}$; in general, we shall set $\neg\varphi \equiv \varphi \to \bot$,

[1] We are not concerned here with the broader debate on the logic of vagueness. If we were, degrees of truth and truth-functionality themselves would need defence, let alone the specific choice of a class of functions to interpret connectives. For an essay on the subject that is sympathetic with fuzzy logic, the interested reader is referred to [11].

[2] Throughout, 't-norm' means '*continuous* t-norm'; we never consider non-continuous t-norms in this paper.

but then $\neg$ may not satisfy the Double Negation Law $\neg\neg\varphi \leftrightarrow \varphi$. (As usual, $\leftrightarrow$ is a shorthand for $(\varphi \to \psi) \wedge (\psi \to \varphi)$.) Similarly, we may not be able to syntactically define a disjunction $\vee$ interpreted by $+^*$. This breaks the classical symmetry between conjunctions and disjunctions in favour of the latter. To see why such a choice is not capricious, one needs to focus on the importance of implication. Of course, one can interpret $\vee$ by a t-conorm. But then, crucially, there will be no relationship between disjunction and implication akin to (**). Instead, in classical logic there is an analogous relationship between $\vee$ and the *dual* of implication, namely, *logical difference*: $\varphi \ominus \psi \equiv \varphi \wedge \neg\psi$. Traditionally, though, $\ominus$ has far less importance than $\to$; the latter is intimately related to inference, whereas the former is not. For instance, in intuitionistic logic one has precisely the same residuation relationship between conjunction and implication as in the classical case, while the classical symmetry between disjunction and conjunction is lost. Nonetheless, as we shall see below, in some important cases the t-conorm does turn out to be syntactically definable. It should be noted that this does not necessarily require $\neg$, as introduced above, to satisfy the Double Negation Law; cf. e.g. the case of Gödel logic. By contrast, it can be proved [5, p. 35–36] that – whichever the t-norm $*$ is – it is possible to reobtain Zadeh's original [12] conjunction and disjunction operators $\min$ and $\max$ in terms of $*$ and $\to^*$, respectively. Hence, *min-conjunction* and *max-disjunction*, as we shall call them, are always available in the present setting.

To sum up, the choice of a specific t-norm $*$ determines a $[0,1]$-valued semantics for the propositional logic $\mathscr{L}$, with conjunction and implication interpreted via $*$ and $\to^*$, respectively. On the semantic side, the definition of $\mathscr{L}$ is completed by stipulating that the tautologies of $\mathscr{L}$ are again defined as those formulas that evaluate to 1 in every possible world.

This framework encompasses as special cases long-studied many-valued logics such as *Łukasiewicz* and *Gödel logic*. These are determined, respectively, by the *Łukasiewicz t-norm* $\max\{0, x+y-1\}$, and its residual *Łukasiewicz implication* $\min\{1, y+1-x\}$; and by the *Gödel* (*minimum*) *t-norm* $\min\{x,y\}$, and its residual

$$x \to^{\min} y = \begin{cases} 1 & \text{if } x \le y, \\ y & \text{otherwise.} \end{cases}$$

Negation (defined by $\neg\varphi \equiv \varphi \to \bot$, as mentioned above) is then interpreted by the involution $1-x$ in Łukasiewicz logic, and by the function

$$f_\neg(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

in Gödel logic. Further, in both cases a syntactic counterpart of the appropriate dual t-conorm is available. For the Łukasiewicz t-norm, the dual t-conorm is $\max\{1, x+y\}$, and its syntactic counterpart can be defined by the De Morgan Laws; for the Gödel t-norm, the dual t-conorm is $\max\{x,y\}$, and its syntactic counterpart can be defined as $\varphi \vee \psi \equiv ((\varphi \to \psi) \to \psi) \wedge ((\psi \to \varphi) \to \varphi)$.

For both Łukasiewicz and Gödel logic, a completeness theorem is available. This uses appropriate finite sets of axiom schemata, and again *modus ponens* as the only deduction rule. For Gödel logic, one also has completeness for arbitrary theories; for Łukasiewicz logic, completeness holds for finitely axiomatisable theories too, but may fail for arbitrary theories (see [7, Thm. 4.6.6]). We stress that these completeness theorems syntactically capture the same notion of tautology (or, more generally, semantic consequence) as in the classical case – formulas satisfied to degree 1 in every possible world. In other words, although the logics at hand are $[0,1]$-valued, the inferential mechanism used only allows deduction of absolutely true propositions from absolutely true assumptions. For more details, please refer to [5].

III. A Logical Analysis of Mamdani-type Fuzzy Control Systems

A. Fuzzification

We consider (physical) observables $x, y, w, \ldots$, which we identify with their values, as measured on the normalised[3] scale $[0,1]$. We write $x(t) \in [0,1]$ for the value of x measured at time t. With each such observable $x, y, w, \ldots$, we associate a finite set of atomic sentences

$$X_1, X_2, \ldots\ ,\ Y_1, Y_2 \ldots\ ,\ W_1, W_2, \ldots \tag{1}$$

The *language* V associated to the (control) system is then the union of the atomic sentences (1). Consider now the set of formulas $\textsc{Form}_\mathrm{V}$ over V. By a *fuzzy set* over the observable x we just mean a function $f\colon [0,1] \to [0,1]$. By the *fuzzification* of the observables $x, y, w, \ldots$ we mean the choice of a finite number of fuzzy sets over each observable; this yields functions

$$f_1, f_2, \ldots\ ,\ g_1, g_2 \ldots\ ,\ h_1, h_2, \ldots\ . \tag{2}$$

In respect of logic, the rôle of f_1, say, is to provide a means of assigning a *degree of truth* (or *truth value*) to the corresponding atomic sentence X_1. In detail, suppose at time t we observe x to be $x(t) \in [0,1]$. Then the degree of truth of X_1 is $f_1(x(t))$. It follows at once that the fuzzification of the observables $x, y, w, \ldots$ uniquely determines at each instant t an atomic assignment (i.e. an assignment to atomic sentences) $\bar\mu\colon \mathrm{V} \to [0,1]$ by

$$\bar\mu_t(X_i) = f_i(x(t))\ ,\ \bar\mu_t(Y_i) = g_i(y(t))\ , \ldots \tag{3}$$

Up to this point, it is immaterial which specific logic we choose to model the system. We now need to make that choice. Fix a many-valued logic $\mathscr{L}$ as in Sect. II, by choosing a t-norm $*\colon [0,1]^2 \to [0,1]$. Then, since $\mathscr{L}$ is truth-functional, the atomic assignment (3) uniquely extends to an assignment

$$\mu_t\colon \textsc{Form}_\mathrm{V} \to [0,1]\ . \tag{4}$$

[3]In the present theoretical context, we gloss over the (often significant) issues involved in normalisation for actual applications.

To sum up, given the fuzzification (2) and the logic $\mathscr{L}$, at each instant t we are able to assign a unique truth value to each formula $\varphi \in \mathrm{FORM}_{\mathrm{V}}$ of $\mathscr{L}$.

Remark: Given the approach we are following in this paper, a relevant issue one may raise at this point is to clarify the rôle of fuzzy sets in purely logical terms. Since the main focus of this paper is the logical interpretation of fuzzy *inference*, due to space limitations we are unable to tackle that issue here. However, we do mention in passing that the chosen fuzzy sets in fact provide an encoding of a theory over the base logic $\mathscr{L}$; and that techniques are available – possibly under appropriate assumptions – to afford the extraction of that theory as an explicit axiomatisation. In the case of Gödel logic, for families of fuzzy sets that form a Ruspini partition, see the results obtained in [13]; cf. also [14].

B. Fuzzy Inference

We now come to the core of fuzzy control systems, namely, fuzzy inference. We discuss Mamdani-type inference first, and then turn to our own recasting in logical terms.

1) Mamdani-type Fuzzy Inference: Our account here will not match Mamdani's original paper [1] exactly; it will rather aim at being consistent with current common practice. Indeed, it may be read as a summary of the implementation of Mamdani-type inference provided by MATLAB's Fuzzy Logic Toolbox. This being understood, fuzzy inference may be summarised as follows.

(M1) Partition the observables $x, y, w, \ldots$ into two disjoint, non-empty subsets, the *input observables* $\{x, y \ldots\}$, and the *output* (*controlled*) *observables* $\{w, \ldots\}$. Partition the language V accordingly, into the sets of *input variables* $\{X_1, X_2, \ldots, Y_1, Y_2, \ldots\}$, and *output variables* $\{W_1, W_2, \ldots\}$. Also partition the fuzzy sets accordingly, into the collections of *input fuzzy sets* $\{f_1, f_2, \ldots, g_1, g_2, \ldots\}$, and *output fuzzy sets* $\{h_1, h_2, \ldots\}$.

(M2) Fix a finite set of *rules* of the form:

`IF` x `is (NOT)` X_i `AND` y `is (NOT)` Y_j `AND ...`
`THEN` w `is (NOT)` W_k `.`

When in 'x `is (NOT)` X_i' the `NOT` is included, we say X_i is *negated*, and similarly for $Y_j, W_k, \ldots$.

(M3) Given the observed input values $x(t), y(t), \ldots$ at time t, for each rule R as in (M2), set

$$R_t = \min\{f'_i(x(t)), g'_j(y(t)), \ldots\} \in [0,1] \ ,$$

where

$$f'_i(x(t)) = \begin{cases} 1 - f_i(x(t)) & \text{if } X_i \text{ is negated in } R; \\ f_i(x(t)) & \text{otherwise.} \end{cases}$$

and similarly for $g'_j(y(t)), \ldots$.

(M4) For each R_t computed in (M3), define the *output fuzzy set of rule R* (*at time t*) as the function $h^t_k \colon [0,1] \to [0,1]$ given by

$$h^t_k(w) = \begin{cases} \min\{R_t, 1 - h_k(w)\} & \text{if } W_k \text{ is negated} \\ & \text{in } R; \\ \min\{R_t, h_k(w)\} & \text{otherwise.} \end{cases}$$

(M5) Define the (*aggregate*) *output fuzzy set* (*at time t*) as the function $F_t \colon [0,1] \to [0,1]$ given by

$$F_t = \max\{h^t_k\} \, ,$$

where the maximum (computed pointwise) ranges over the output fuzzy sets, as in (M4), of all rules.

Thus, at each instant t, Mamdani-type inference produces as a result the fuzzy set F_t as in (M5).

Remark: Several variants of the procedure described above are used in practice. Let us mention some common ones.

- One can allow an `OR` connective in the rules' antecedents along with `AND`, interpreting it by a suitable operator such as the maximum or some other t-conorm.
- One can use a different operator than the minimum – e.g. a different t-norm – to interpret the `AND` connective.
- One can use a different operator than the maximum – e.g. a different t-conorm – to compute the aggregate fuzzy set.
- Somewhat less frequently, one can use a different operator than the minimum to interpret the `IF...THEN...` connective in computing the output fuzzy set of each rule.
- Applications generally need to deal with more than one output observable. Here we chose to describe the case with one observable only, to avoid a heavier notation; the constraint is immaterial for our present purposes.

2) Logic-based Alternative to Mamdani-type Inference: We describe our own procedure to obtain an "output fuzzy set" given values for the "input observables". Extended comments on its logical meaning, and on its relationship to Mamdani-type inference are provided in Sect. IV and V.

(L0) Fix a t-norm $*$, and its corresponding logic $\mathscr{L}$ (cf. Sect. II).

(L1) Same as (M1).

(L2) Fix a theory Θ in the language V. The most important case in applications is when V is finite, and Θ is finitely axiomatisable; we assume this throughout. Then Θ can be safely thought of as a finite set of formulas $\theta_1, \ldots, \theta_m$ in the finite set of variables $\mathrm{V} = \{X_1, \ldots, Y_1, \ldots, W_1, \ldots\}$, or even as the single formula

$$\Theta = \theta_1 \wedge \cdots \wedge \theta_m \, , \tag{†}$$

where $\wedge$ is the conjunction of $\mathscr{L}$, semantically interpreted by $*$. In the sequel, we work with Θ as in (†).

(L3) Define the function T_Θ as the *truth-value function* of Θ. That is, given $x, y, w, \ldots \in [0,1]$, consider the unique assignment $\mu \colon \mathrm{FORM}_{\mathrm{V}} \to [0,1]$ such that $\mu(X_i) = f_i(x)$, $\mu(Y_j) = f_j(y)$, $\mu(W_k) = h_k(w)$, $\ldots$. Then

$$T_\Theta(x, y, \ldots, w, \ldots) = \mu(\Theta) \, .$$

(L4) Given the observed values $x(t), y(t), \ldots$ at time t, define the *output truth-value function*

$$T_t \colon [0,1]^n \to [0,1] \ ,$$

where n is the cardinality of the set of output variables $\{W_1, \ldots\}$, by

$$T_t(w, \ldots) = T_\Theta(x(t), y(t), \ldots, w, \ldots) .$$

Here, note that $x(t), y(t), \ldots$ are constants, whereas $w, \ldots$ are not.

(L5) For each instant t, set

$$M_t = \{p \in [0,1]^n \mid T_t(p) \text{ is maximal in the range of } T_t\} ;$$

we call M_t the set of *maximising output values*.

Thus, at each instant t, the procedure above yields as a result the set of maximising output values M_t as in (L5).

C. Defuzzification

In light of (M5) above, the outcome of a Mamdani-type inference at time t is the fuzzy set $F_t \colon [0,1] \to [0,1]$. The domain of F_t is the normalised range of values of the physical observable w to be controlled. For actual control to be achieved, one must set w to a specific value $w_t \in [0,1]$. This is effected via *defuzzification of* F_t. Throughout this paper and its sequel, we assume that w_t is computed from F_t using the well-known *mean of maxima* defuzzification method. With an eye to the second, experimental part of this piece of work [6], let us merely indicate how one computes w_t when using a discrete approximation to F_t. Choose an integer $N \geq 1$, and consider the set of sample points $S = \{\frac{1}{n} \mid n = 0, 1, \ldots, N\} \subseteq [0,1]$. From these, extract those that maximise F_t over S; i.e., set $M_t = \{s \in S \mid F_t(s) = \max_{w \in S} F_t(w)\}$. Finally, set

$$w_t = \frac{\sum_{s \in M_t} s}{|M_t|} , \qquad \text{(MOM)}$$

where $|A|$ denotes the cardinality of a set A.

Concerning the procedure described in Sect. III-B.2, in light of (L5) above its outcome is the set M_t of maximising output values. To extract a single value from this set, let us assume that only one output observable w is given, in line with our presentation of Mamdani-type inference. Then we again obtain w_t by computing the average in (MOM).

In summary, from the point of view of defuzzification the two procedures considered in Sect. III-B are in agreement – they both use the mean of maxima method.

Remark: While defuzzification is an important topic with a substantial literature, here we are not concerned with it, where the focus is on inference. A relevant research problem along the lines discussed in this paper is to clarify the rôle of defuzzification in respect of logic. Although this is the subject of current investigation, let us mention in passing the paper [15], where the connection between defuzzification methods and the expected truth value of formulas (with respect to a measure on the set of possible worlds) is discussed for Gödel logic.

IV. Discussion and Remarks

We begin discussing the basic ideas behind the procedure in Sect. III-B.2. The theory Θ provides a linguistic description of relationships[4] among the atomic sentences $X_i, Y_j, W_k, \ldots$. In general, it is customary to think of Θ as reflecting an agent's knowledge about the meaning – more precisely, about the *intended interpretation* – of such sentences in the actual world. For instance, if the intended interpretation of X_1 is "*It is cold in this room*", and that of X_2 is "*It is warm in this room*", then the agent's knowledge about the meaning of the atomic propositions may include the fact that $\neg(X_1 \vee \neg X_2)$ is always true.[5] Formally, this precisely amounts to saying that the formula $\neg(X_1 \vee \neg X_2)$ is part of the theory Θ.

However, in certain contexts it is appropriate to think of Θ as more than just a formal model of an agent's knowledge. For if our aim is, say, to achieve control of a given physical system, then Θ should be thought of as a summary of our knowledge about the system, along with a description of the *desirable states* of the system. Continuing our previous example, suppose we are attempting to keep the temperature of the room within a certain range, so that it is not too hot nor is it too cold, by (automatically) operating the heating system in the room. Then Θ might also include – besides the previous formula $\neg(X_1 \vee \neg X_2)$, and among others – a formula modelling the sentence:

"*If this room is cold, then the heating system is on.*"

For this, we prepare an additional propositional variable W_1, whose intended interpretation is "*The heating system is on*", and we add the formula $X_1 \to W_1$ to our theory Θ. There are three key points to note about $X_1 \to W_1$ in this example. First:

(P1) *It is certainly possible that, in the actual world, $X_1 \to W_1$ is false.*[6]

For it is perfectly conceivable that the room is cold, and the heating system is turned off. Hence $X_1 \to W_1$ does *not* embody knowledge about the real world in the same sense as $\neg(X_1 \vee \neg X_2)$ did. Further:

(P2) *It is likewise possible that, in the actual world, $X_1 \to W_1$ is true, and that $\neg(X_1 \to W_1)$ fails to be true.*

For it is perfectly conceivable that the room is cold, and the heating system is turned on. Thus, (P1) and (P2) tell us that there are *possible worlds* – i.e. sufficiently specified situations – that are consistent with the truth of $X_1 \to$

[4] Strictly speaking, such relations are encoded by Θ, *along with the chosen fuzzification*, as mentioned in the Remark at the end of Sect. III-A.

[5] Of course, if the agent in question is a many-valued reasoner, she may not necessarily want to assume that much, depending on her own reading of "*cold*" and "*warm*", i.e. on her specific choice of the logic $\mathscr{L}$ used to model those notions and their mutual relationships.

[6] Or false to some degree, in the many-valued setting. The general point we are making in this discussion does not depend in an essential manner on the fact that the underlying logic $\mathscr{L}$ be many-valued, nor does it depend on the specific semantics of connectives.

W_1, and others that are consistent with its falsehood. This notwithstanding, we have:

(P3) *In order to achieve control in our present example, it is reasonable to include $X_1 \to W_1$ in our theory Θ, whereas it would not be reasonable to include its negation $\neg(X_1 \to W_1)$.*

Indeed, let us assume for the sake of argument that the underlying logic $\mathscr{L}$ is Boolean, and let us see how $X_1 \to W_1$ may be used to achieve (partial) control of the room temperature, following the idea behind the procedure in Sect. III-B.2. The crux of (L0–L5) is simple to state:

Given the truth values of the input propositional variables $X_i, Y_j, \ldots$, choose truth values for the output propositional variables $W_k, \ldots$, that make the overall truth value of Θ as large as possible.

In our example, then, what we need to do is simply trying to make $X_1 \to W_1$ as true as possible – i.e. just true, in classical logic – by *choosing the truth value of W_1, given the truth value of X_1*. So if we start from a situation where X_1 is true, we have no choice for the truth value of W_1 – it must be 1. Of course, in *operational terms*, this amounts to saying that whenever the room is cold, we must turn the heating system on. On the other hand, if we replaced $(X_1 \to W_1)$ by $\neg(X_1 \to W_1)$, then we would leave the heating system off even if the room is cold. While that situation, as discussed, corresponds to a perfectly legitimate possible world, it is one that is not desirable for our present aims. In the language of control systems, we can sum up our discussion so far as follows.

The theory Θ is to be thought of as a description of those states in which the system is controlled, i.e. of the desirable states *of the system.*

Insofar as Θ provides a *good* description of the desirable states, one can use it as a *recipe to control the system*. That recipe is what steps (L0–L5) in Sect. III-B.2 spell out in detail. We now remark on some of the main differences between (L0–L5) and (M1–M5), though it is in fact the similarities that we regard as more important, a point we shall return to these in the next section.

- The syntactic restrictions in (M2) are guided by the background assumption that one is aiming at approximating a function, namely, the control function of the system. By contrast, in (L2) we put no restriction at all on Θ. Indeed, in contexts where one uses (L0–L5) not to control a system, but, say, to take a decision on the basis of vague information, it would not be appropriate to restrict Θ to implicative formulas. Additional research may lead to reasonable syntactic constraints for Θ under various assumptions.
- The procedure in (L0–L5) is fully general in another respect, namely, the choice of the logic $\mathscr{L}$. The latter can be any t-norm-based logic as in Sect. II, provided only $\mathscr{L}$ enjoys completeness for (finitely axiomatisable) theories with respect to its $[0, 1]$-valued semantics. Although we cannot discuss the significance of this completeness requirement in full here, we provide the following informal comment, by way of illustration of some of the issues involved. In the absence of completeness, there may be *unobservable possible worlds described by* Θ. That is, Θ describes situations which can actually take place (assuming Θ is a faithful linguistic description of our intended models), but that do not correspond to *any* choice of values $x_t, y_t, w_t \ldots \in [0, 1]$ of the physical observables. That happens, for instance, if $\mathscr{L}$ is Łukasiewicz logic, and Θ is a so-called *non-semisimple* theory; in this case, it can be shown that Θ is necessarily not finitely axiomatisable. Details on this phenomenon can be found in [7].
- In the two procedures, (M5) and (L4) are analogous steps. That is, the output aggregate fuzzy set F_t in the former corresponds, conceptually, to the output truth-value function T_t in the latter. But whereas Mamdani-type inference has F_t as its end result, our own version does not. We take one more step, and compute the set of maximising output values M_t in (L5). We do so because the logical bases of our procedure make (L5) a completely justified step – more than that, an unavoidable one. Indeed, observe that M_t determines the collection of possible worlds wherein Θ is satisfied to maximal degree, given the truth values of the input propositional variables $X_i, Y_j, \ldots$. Now if we take seriously the claims that (i) $\mathscr{L}$ is a logic of vague propositions suited to our problem, so that larger truth values of propositions indicate statements that are closer to absolute truth, and that (ii) Θ expresses, in the language of $\mathscr{L}$, sound (possibly partial) knowledge about the problem at hand and the desirable states of the system, then no other conclusion is possible than the following one:

  The more we can make Θ closer to truth in the actual world (*by acting on the truth values of the output variables $W_k, \ldots$*), *the closer we are to attaining a desirable state of the system.*

  Settling for situations where Θ is satisfied to *lower* degrees cannot possibly improve the situation, unless we reject either (i) or (ii). Hence, if the end result of the procedure is unsatisfactory in a specific case, the problem must lie with $\mathscr{L}$ or with Θ, not with (L5).
- By contrast, there is no way that Θ can help us in selecting a specific element of M_t – which is why the output of our procedure must be the whole of M_t. Indeed:

  Two distinct elements of the set M_t in (L5) *are entirely indiscernible on the grounds of the knowledge embodied by Θ.*

  In other words, Θ cannot help us at all with defuzzification.
- The combined effect of (M1–M5) and (MOM) method is equivalent to the combined effect of (L0–L5) and (MOM) for a specific choice of $\mathscr{L}$. This will be discussed in the next section. It should nonetheless be

observed that the two procedures differ at a fundamental level in their views of defuzzification. Mamdani-type systems adhere to the traditional fuzzy view that the whole information carried by the output fuzzy set may be used to obtain a crisp output value. On the other hand, our procedure insists that it is *always* appropriate to consider maximising values of the output truth-value function, for the reasons expounded above. The maximisation process is therefore to be seen as an integral part of the procedure.

V. Recovering Mamdani-type Fuzzy Inference

In this short section we indicate how to recover the Mamdani-type fuzzy inference (M1–M5) described in Sect. III-B.1 as a special case of our procedure (L0–L5). To this end, we let $\mathscr{L}$ be Łukasiewicz logic. Recall that $\neg$ is then interpreted by $f_{\neg}(x) = 1 - x$, and that min-conjunction and max-disjunction are definable. In the rest of this section, we denote the latter by $\wedge$ and $\vee$, respectively. The remaining connectives of Łukasiewicz logic will play no rôle for our purposes here. Given now the fuzzy sets for the Mamdani-type system

To see that with the assumptions above (L0–L5) reduce to (M1–M5), consider by way of example the two rules

$$\texttt{IF } x \texttt{ is } X_1 \texttt{ AND } y \texttt{ is } Y_1 \texttt{ THEN } w \texttt{ is } W_1\,. \tag{R1}$$

$$\texttt{IF } x \texttt{ is } X_2 \texttt{ THEN } w \texttt{ is NOT } W_2\,. \tag{R1}$$

Translate (R1) and (R2) into a formula of $\mathscr{L}$ as follows:

$$\theta = (X_1 \wedge Y_1 \wedge W_1) \vee (X_2 \wedge \neg W_2)\,.$$

Let Θ be the theory over $\mathscr{L}$ axiomatised by θ; cf. with (†) in (L2). Now fix an instant of time t, along with x_t, $y_t \in [0,1]$. To these observed values correspond truth values $f_1(x_t), f_2(x_t), g_1(y_t)$ of X_1, X_2, and Y_1, respectively. Set $m = \min\{f_1(x_t), f_2(x_t), g_1(y_t)\}$. Then the output truth-value function T_t in (L4) is given by

$$\max(\min(f_1(x_t), g_1(y_t), h_1(w)), \min(f_2(x_t), 1 - h_2(w))).$$

Direct inspection shows that the output fuzzy set $F_t(w)$ of (M1–M5) coincides with $T_t(w)$, as was to be shown.

It is clear that the example we considered here is readily extended to cover the general case.

Acknowledgment

This work was supported in part by the Princes Gate Group under research contract no. 11103 with the Università degli Studi di Milano.

References

[1] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1–13, 1975.

[2] L. A. Zadeh, "Outline of new approach to the analysis of complex systems and decision processes," *IEEE Trans. Systems, Man, and Cybernet.*, vol. SMC-3, pp. 28–44, 1973.

[3] C. C. Lee, "Fuzzy logic in control systems: fuzzy logic controller. I, II," *IEEE Trans. Systems Man Cybernet.*, vol. 20, no. 2, pp. 404–418, 419–435, 1990.

[4] A. Sala, T. M. Guerra, and R. Babuška, "Perspectives of fuzzy systems and control," *Fuzzy Sets and Systems*, vol. 156, no. 3, pp. 432–444, 2005.

[5] P. Hájek, *Metamathematics of Fuzzy Logic*. Dordrecht: Kluwer, 1998.

[6] S. Bova, P. Codara, D. Maccari, and V. Marra, "A Logical Analysis of Mamdani-type Fuzzy Inference, II. An Experiment on the Technical Analysis of Financial Markets," *FUZZ-IEEE 2010. IEEE International Conference on Fuzzy Systems*, 2010, submitted.

[7] R. L. O. Cignoli, I. M. L. D'Ottaviano, and D. Mundici, *Algebraic foundations of many-valued reasoning*, ser. Trends in Logic—Studia Logica Library. Dordrecht: Kluwer Academic Publishers, 2000, vol. 7.

[8] R. Keefe, *Theories of Vagueness*. Cambridge: Cambridge University Press, 2000.

[9] V. Novák, "Linguistically oriented fuzzy logic control and its design," *Internat. J. Approx. Reason.*, vol. 12, no. 3-4, pp. 263–277, 1995.

[10] F. Klawonn and V. Novák, "The relation between inference and interpolation in the framework of fuzzy systems," *Fuzzy Sets and Systems*, vol. 81, no. 3, pp. 331–354, 1996.

[11] N. J. Smith, *Vagueness and degrees of truth*. New York: Oxford University Press, 2008.

[12] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[13] P. Codara, O. M. D'Antona, and V. Marra, "An analysis of Ruspini partitions in Gödel logic," *Internat. J. Approx. Reason.*, vol. 50, no. 6, pp. 825–836, 2009.

[14] ——, "A characterisation of bases of triangular fuzzy sets," *FUZZ-IEEE 2009. IEEE International Conference on Fuzzy Systems*, pp. 604–609, 2009.

[15] S. Aguzzoli, B. Gerla, and V. Marra, "Defuzzifying formulas in Gödel logic through finitely additive measures," *FUZZ-IEEE 2008. IEEE International Conference on Fuzzy Systems*, pp. 1886–1893, 2008.