Conditional Constraints, Implication Based Rules, and Possibilistic Rule Bases: Are They Any Good?

Joachim Weisbrod
Forschungszentrum Informatik
Karlsruhe
weisbrod@fzi.de

Martin Spott
IPD
Universität Karlsruhe
spott@ipd.info.uni-karlsruhe.de

Abstract — To answer the question imposed in the title right away: Yes, they are. Conditional constraints, implication based rules, and possibilistic rule bases are different notions representing the very same concept. And this concept is as useful and interesting as the well known and widely used Mamdani approach. So far, real success in applying fuzzy rule bases is restricted to fuzzy control and Mamdani inference. We claim that the lack of applications making use of possibilistic reasoning is mainly due to a lack of understanding how to deal with rules based on possibility distributions.

Mamdani inference as used in fuzzy control and possibilistic reasoning are complementary mechanisms. They are complementary with respect to the way they deal with incomplete and inconsistent information. From that point of view, it is not surprising that using the very same rule base in both settings does not work. This means, that the standard way of specifying Mamdani knowledge bases does not help in the possibilistic case. Based on the assumption that rule based systems are helpful and manageable as long as rules represent local information we present a new way how to specify possibilistic rule bases. In order to prove the usefulness of this approach, we show that there are some major drawbacks and limitations of Mamdani inference, that are easily solved by correctly applying possibilistic reasoning.

We do not claim, however, that possibilistic reasoning is the one and only mechanism to choose. Both mechanisms have their advantages and drawbacks and it heavily depends on the problem at hand whether one or the other mechanism should be chosen. In complex situations we expect a combined mechanism to be useful.

I. Introduction

The application of fuzzy rules has always been one of the major concerns within the fuzzy research community. After some years of general confusion generated by the vast amount of possibilities, today opinions seem to converge. From out point of view, there are two fundamental ways how to deal with fuzzy rules.

The first one is the well known Mamdani approach, that has been proven to be a very successful paradigm in lots of industrial applications. There are several ways how to look at Mamdani inference. ZADEH calls it joint constraint [1], DUBOIS & PRADE call it conjunction based [2], KRUSE calls it equality based [3]. These different naming conventions reflect the fact, that some years ago lots of researchers tried to answer the question, why Mamdani inference is based on conjunction rather than implication. And in order to come up with a satisfying answer, several theoretical skeletons for the Mamdani approach have been introduced. Today, Mamdani inference is mostly interpreted as some sort of fuzzy interpolation.

The second basic inference mechanism is the possibilistic one. It is based on the notion of a possibility distribution. From the semantics of possibility, this inference mechanism based on the Gödel implication relation may be derived. The possibilistic approach to fuzzy reasoning is also called conditional constraint or implication based.

In recent work, the authors showed that both inference mechanisms are very strongly related to each other. We showed that Mamdani inference is based on the notion of a support distribution in the very same way that possibilistic reasoning is based on possibility distributions. Being based on either support distributions $\sigma$ or possibility distributions $\pi$, we call the mechanisms $\sigma$-reasoning and $\pi$-reasoning, respectively.

As pointed out in [4], [5], $\sigma$-reasoning and $\pi$-reasoning possess complementary advantages and drawbacks when it comes to dealing with fuzzy -i.e. imperfect- knowledge. The possibilistic approach is appropriate for dealing with incomplete knowledge, but sensitive to contradictions. For $\sigma$-reasoning, the opposite holds. Despite its well understood theoretical background, possibilistic reasoning has not gained much practical attention so far. Today, practical success is restricted to $\sigma$-reasoning as originally developed by MAMDANI [6]. From our point of view, this is mainly due to the fact that so far there has been no correct understanding how to specify possibilistic rule bases.

In this paper we argue, that this complete reliance on $\sigma$-reasoning as soon as practical interests are involved, should change. We do not claim that $\pi$-reasoning is the better mechanism. But we will show that Mamdani inference has some limitations where possibilistic mechanisms can help. In section II we will present some basics, in order to identify the fundamental differences between $\sigma$-reasoning and $\pi$-reasoning. The main limitation of Mamdani inference is addressed in section III. Section IV deals with the consequences of these considerations with respect to fuzzy rule bases. Here we present the main result of this paper, namely how to specify possibilistic rule bases. In section V, some experiments are presented that support the statements and claims made before.

II. Support and Possibility Distributions

Assume we are dealing with a variable $x$. We know that $x$ holds a value $u' \in \mathcal{U}_x$. Nevertheless, in the general case we do not know this value $u'$ exactly, but have only some more or less fuzzy knowledge about this actual value $u'$.
As introduced by Zadeh [7], a possibility distribution \( \pi_x : \mathcal{U}_x \rightarrow [0, 1] \) represents an elastic constraint on the variable \( x \) that may be used in order to specify such a kind of knowledge. The value \( \pi_x(u) \) represents the degree of possibility that the assumption \( x = u \) is true. If we are completely ignorant about \( x \), \( \pi_x \equiv 1 \) holds. The more we know about \( x \), the less values \( u \in \mathcal{U}_x \) are considered to be possible. This is, the more information we get, the smaller the corresponding possibility distribution gets. Consequently, given two elastic constraints \( \pi_x \) and \( \pi'_x \) on \( x \), the resulting constraint is determined by the conjunction \( \pi_x \cap \pi'_x \) of the original distributions. That is, possibility distributions are aggregated by means of conjunction.

A support distribution \( \sigma_x : \mathcal{U}_x \rightarrow [0,1] \), on the other hand, represents an elastic hypothesis on the variable \( x \). Now the value \( \sigma_x(u) \) represents the degree of support that the assumption \( x = u \) is true. In this setting, complete ignorance about \( x \) is represented by \( \sigma_x \equiv 0 [8], [9] \). And the more we know about \( x \), the more values \( u \in \mathcal{U}_x \) may get support. With additional information, support can only increase. In this setting, distributions are aggregated by means of disjunction. Given two pieces of information \( \sigma_x \) and \( \sigma'_x \) on the actual value of \( x \), we combine these by determining \( \sigma_x \cup \sigma'_x \).

As subtle this difference seems at first glance, as fundamental are its consequences. In the first case, impossibility, i.e. \( \pi_x(u) = 0 \), is real information. The smaller the possibility gets, the more important is the information. In the second case, full support \( \sigma_x(u) = 1 \) is significant information, and small values are less interesting. Possibility distributions represent negative knowledge, where assumptions are more or less excluded. Support distributions represent positive knowledge, where assumptions are more or less endorsed.

If we look at a fuzzy set \( \tilde{A} \) like the one depicted on the upper left of figure 2, it becomes obvious that it makes a fundamental difference whether we interpret the fuzzy predicate “\( y \) is \( \tilde{A} \)” as a possibility distribution \( \pi_y \equiv \mu_A \) or as a support distribution \( \sigma_y \equiv \mu_A \). Considered as a possibility distribution, the predicate “\( y \) is \( \tilde{A} \)” is a very tough statement: most of the values \( v \in \mathcal{U}_y \) are declared impossible, and the remaining ones are only possible to some degree. If considered as a support distribution, on the other hand, the very same predicate is a very cautious statement, as only few elements of \( \mathcal{U}_y \) are slightly supported.

Unsurprisingly, different semantics imply different mechanisms when it comes to applying fuzzy rule bases. We may look at the fuzzy rule “IF \( x \) is \( \tilde{A} \) THEN \( y \) is \( \tilde{B} \)” (or \( [\tilde{A}_x \Rightarrow \tilde{B}_y] \), for short) from two different points of view: If we consider the fuzzy predicates “\( x \) is \( \tilde{A} \)” and “\( y \) is \( \tilde{B} \)” as being support distributions \( \sigma_x \equiv \mu_A \) and \( \sigma_y \equiv \mu_B \), respectively, the relation induced between \( x \) and \( y \) is represented by the well known Mamdani relation \( \text{Mamdani}[\tilde{A}_x \Rightarrow \tilde{B}_y] \) [8], [9]:

\[
\sigma_{x,y}(u,v) := \text{Mamdani}[\tilde{A}_x \Rightarrow \tilde{B}_y](u,v) \quad (1)
\]

\[
= \min \{ \mu_A(u), \mu_B(v) \} \quad (2)
\]

\[
= \min \{ \sigma_x(u), \sigma_y(v) \}. \quad (3)
\]

If, on the other hand, we look at the predicates from a possibilistic point of view, the one and only implication relation to choose is the so called G"odel implication relation \( \text{G"odel}[\tilde{A}_x \Rightarrow \tilde{B}_y] \), defined by

\[
\pi_{x,y}(u,v) := \text{G"odel}[\tilde{A}_x \Rightarrow \tilde{B}_y](u,v) \quad (4)
\]

\[
= \begin{cases} 
1 & \text{if } \mu_A(u) \leq \mu_B(v), \\
\mu_B(v) & \text{else}. 
\end{cases} \quad (5)
\]

\[
= \begin{cases} 
1 & \text{if } \pi_x(u) \leq \pi_y(v), \\
\pi_y(v) & \text{else}. 
\end{cases} \quad (6)
\]

Given several rules \( [\tilde{A}_i \Rightarrow \tilde{B}_i] \), the corresponding joint distributions \( \pi_{x,y} \equiv \text{Mamdani}[\tilde{A}_i \Rightarrow \tilde{B}_i] \) and \( \pi_{x,y} \equiv \text{G"odel}[\tilde{A}_i \Rightarrow \tilde{B}_i] \) have to be aggregated by disjunction and conjunction, respectively:

\[
\sigma_{x,y}(u,v) := \bigcup_i \sigma_{x,y}^i \quad (7)
\]

\[
\pi_{x,y}(u,v) := \bigcap_i \pi_{x,y}^i \quad (8)
\]

Finally, in order to determine the result of applying the input “\( x \) is \( \tilde{A} \)” to such a rule base, in both settings we make use of max-min composition \( \sigma_{x,y} \equiv \mu_{\tilde{B}} \), we compute \( \sigma_{x,y} \cap \sigma_{x,y} \) and \( \sigma_{x,y} \cap \sigma_{x,y} \), defined as

\[
\sigma_y(v) := \max_{u \in \mathcal{U}_x} \min \{ \sigma_y^u(u), \sigma_{x,y}(u,v) \} \quad (9)
\]

and

\[
\pi_y(v) := \max_{u \in \mathcal{U}_x} \min \{ \pi_y^u(u), \pi_{x,y}(u,v) \}. \quad (10)
\]

For a theoretic derivation and a thorough analysis of these results see [8], [9]. We refer to the complete inference mechanism based on support distributions as \( \sigma \)-reasoning or Mamdani inference. Consequently, the possibilistic scheme is called \( \pi \)-reasoning, G"odel inference, or simply possibilistic inference.

III. LIMITATIONS OF MAMDANI INFERENCE

The main problem of dealing with positive knowledge represented by support distributions is related to the accumulation of gradual inconsistencies. We are using fuzzy models of complex systems in order to have a simplified description of the state of affairs. But this simplification is not for free. A fuzzy description is always an imperfect one. In the domain of fuzzy rule bases, we have to deal with gradual contradictions and gradual incompleteness. In most situations, several rules with different consequences gradually apply, and in the general case some inputs are better matched than others [10].

As stated before, support distributions deal with inconsistent pieces of information by disjunctively combining them. It is not surprising that this mechanism becomes problematic when too much inconsistencies arise. But exactly this
Example 1: Consider the mapping \( y = f(x) \) with \( f(x) = x \) and \( x, y \in [0, 100] \). We describe the mapping \( f \) by a fuzzy rule base specified by means of support distributions. In order to do so, we roughly describe the interval \([0, 100]\) by the nine fuzzy sets in figure 1. Based on the terms \( \tilde{10}_x, \tilde{20}_x, \ldots, \tilde{90}_x \) (\( v \in \{x, y\} \)), the rule base \( F_\sigma \) to represent \( f \) is straightforward:

\[
\begin{align*}
\text{IF } x \text{ is } \tilde{10}_x \text{ THEN } y \text{ is } \tilde{10}_y \\
\text{IF } x \text{ is } \tilde{20}_x \text{ THEN } y \text{ is } \tilde{20}_y \\
\text{IF } x \text{ is } \tilde{30}_x \text{ THEN } y \text{ is } \tilde{30}_y \\
\vdots \\
\text{IF } x \text{ is } \tilde{90}_x \text{ THEN } y \text{ is } \tilde{90}_y 
\end{align*}
\]

Now we examine how this rule base behaves when used in a chain. This is, we determine

\[
y_n = f^n(x) = f(f(\ldots(f(x))\ldots))
\]

by applying the rule base \( \tilde{F}_\sigma \) to its own output \( n \) times:

\[
\tilde{y}_n = (F_\sigma \circ (F_\sigma \circ \ldots (F_\sigma \circ x)\ldots))
\]

We chose as initial input the singleton \( \tilde{x} = 55 \), frankly admitting that this is an unfair input to this type of rule base. Figure 2 presents the results \( \tilde{y}_1 \) to \( \tilde{y}_6 \). Obviously, for \( n > 6, \tilde{y}_n = \tilde{y}_6 = \tilde{y}_0 \) holds.

This is, after just five inference steps the result is completely useless. Every value \( v \in \mathcal{U}_y \) gets the same degree of support. There is no reasonable decision to be made from such a result.

Choosing different initial inputs, the final results may be somewhat more useful. Figure 3 shows the final result \( \tilde{y}_6 = \tilde{y}_n, \) (\( n > 5 \)) for the initial input \( x = 50 \), for instance. This input perfectly matches the fifth rule. Nevertheless, example 1 clearly addresses the fundamental weakness of \( \sigma \)-reasoning: Due to the accumulation of gradual inconsistencies within each inference step, overall results tend to become very broad. When several layers of rules are involved, \( \sigma \)-reasoning - the very same mechanism that is successful in fuzzy control - generally fails. And this is the very reason, why there are no complex rule bases applied in fuzzy control up to today.

It is therefore natural to look for alternative inference schemes. And with \( \pi \)-reasoning there is another well known mechanism to be analyzed with respect to this property. Nevertheless, so far possibilistic inference has not been exploited in a very successful way. In the following section we argue, that this is due to a complete misconception in the way, possibilistic rule bases have been designed so far.

IV. Specifying Possibilistic Rule Bases

Support and possibility distributions are complementary concepts. \( \sigma \)-reasoning and \( \pi \)-reasoning are dual inference mechanisms with complementary advantages and drawbacks. The further deals with positive, the latter with negative information. It is therefore by no means surprising, that using the very same rule base in both settings makes no sense. In order to apply possibilistic inference, the fuzzy control type of rule base as the one seen in example 1 gen-
In this system, for short. However, if we apply the very same rule base in the pessimistic case, there is a problem: If, for example, the input \((x, y)\) lies right between \((\tilde{A}_y, \tilde{B}_y)\) and \((\tilde{A}_x, \tilde{B}_x)\), the output is restricted both to \(\tilde{A}_y\) and \(\tilde{A}_x\). Since both outputs do not overlap, the result is equal to zero – for the output \(z\), no value is possible any more. Figure 5 shows the results for the input \((\tilde{A}_x, \tilde{A}_y)\) using \(\sigma\)-reasoning (left) and \(\pi\)-reasoning (right).

This is not, what the rule base was intended to imply. So, obviously the rule base works for the Mamdani but not for the possibilistic case. Why this? Let’s have a look at a single fuzzy predicate “\(x\) is \(\tilde{A}_x\)”.

If we interpret the fuzzy predicate “\(x\) is \(\tilde{A}_x\)” as a possibilistic piece of information by assigning \(\pi_x := \mu_{\tilde{A}_x}\), the predicate “\(x\) is \(\tilde{A}_x\)” becomes a very significant statement: It definitely excludes all values \(u \in \mathcal{U}_x\) with \(\mu_{\tilde{A}_x}(u) = 0\). Only the values within the set \(\text{supp}(\tilde{A}_x) := \{ u \mid \mu_{\tilde{A}_x}(u) > 0 \}\) are not completely determined yet. Only within \(\text{supp}(\tilde{A}_x)\), i.e. the interval \([10, 30]\), there is still some information missing. In the Mamdani case, using an interpretation \(\sigma_x := \mu_{\tilde{A}_x}\) based in support distributions, this situation changes. Now, the statement “\(x\) is \(\tilde{A}_x\)” is not so significant any more. The information given is restricted to the set \(\text{supp}(\tilde{A}_x) = [10, 30]\), whereas there is no statement whatsoever on values in \(\mathcal{U}_x \setminus \text{supp}(\tilde{A}_x)\).

And exactly this difference makes the problem: In the Mamdani case based on support distributions, a rule like

\[
\text{IF } x \text{ is } \tilde{A}_x \text{ AND } y \text{ is } \tilde{A}_y \text{ THEN } z \text{ is } \tilde{A}_z
\]

reads: “If there is support that \(x\) is about 20 and that \(y\) is about 30, there is support that \(z\) is about 20”. In the possibilistic case, however, the meaning of the rule is: “If we can exclude that \(x\) is not about 20 and that \(y\) is not about 30, we can exclude that \(z\) is not about 20”. Or, in other words: “If we are sure that \(x\) is about 20 and \(y\) is about 30, then definitely \(z\) is about 20”. So, for the input \((x, y) = (\tilde{A}_x, \tilde{A}_y)\) and the first interpretation we derive that eventually \(z\) is about 20, whereas the second interpretation insists in \(z\) to be about 20, since it is impossible for \(z\) not to be about 20. The interpretation based on support distributions makes the consequence “\(z\) is \(\tilde{A}_z\)” a local statement restricted on \(\text{supp}(\tilde{A}_x)\), whereas the possibilistic interpretation excludes nearly the entire domain \(\mathcal{U}_z\) of \(z\) and thus makes a very restrictive and global statement. And very easily such global statements made by different rules contradict with each other, as seen in the example above.

We claim, that the overall success of rule based systems is very much based on the concept of combining local pieces of knowledge, with each of them being represented by a single rule.

In order to apply this perception to the problem of specifying a reasonable possibilistic rule base, we recapitulate the case of \(\sigma\)-reasoning and Mamdani inference:

**How to specify a Mamdani rule base:**

When based on support distributions, rule bases are sets of rules that map observed input states to proposed output states. In order to build such a rule base, we try to map
In order to mimic this successful and popular mechanism when it comes to building possibilistic rule bases, all we have to do is to make possibilistic consequents local statements, as well. And if—as seen before—“$x$ is $\tilde{A}_x$” is a global statement in the possibilistic setting, “$x$ is NOT $\tilde{A}_x$” is a local one: Since the further definitely excludes most of the values $u \in \mathcal{U}$ (namely, all those contained in $[0, 10] \cup [30, 100]$), the latter imposes only some restrictions within the range $supp(\tilde{A}_x) = [10, 30]$. So the basic idea is very simple: Instead of having consequents of the form “$x$ is $A$”, we have to look for consequents of the form “$x$ is NOT $A$”, where $A$ is a typical fuzzy set as used nowadays in order to represent a linguistic term.

In order to clarify and generalize this simple idea we have a closer look at a simple one input / one output system. First of all, we analyze a single rule “IF $x$ is $\tilde{A}$ THEN $y$ is $\tilde{B}$” with $\tilde{A}$ and $\tilde{B}$ being typical trapezoidal membership functions. Figure 6 presents the rule in the context of support distributions. The well known Mamdani relation, i.e. the conjunctive aggregation of the cylindrical extensions of $\tilde{A}$ and $\tilde{B}$, results in a local “fuzzy point” in $\mathcal{U}_x \times \mathcal{U}_y$. Only within the subset $supp(\tilde{A}) \times supp(\tilde{B})$ of $\mathcal{U}_x \times \mathcal{U}_y$ useful information is given, i.e. support is imposed.

If, however, we attach the very same rule $[\tilde{A}_x \Rightarrow \tilde{B}_y]$ with a possibilistic interpretation, the Gödel implication relation makes a global statement (see figure 7): Within the range of $supp(\tilde{A})$, most values $v \in \mathcal{U}_y$ are declared impossible. Realizing the idea presented above, figure 8 presents the Gödel implication relation for the rule $[\tilde{A}_x \Rightarrow \text{NOT} \tilde{B}_y]$. Here we find the same local behaviour as in figure 6. There is only a “fuzzy point” $supp(\tilde{A}) \times supp(\tilde{B})$ in $\mathcal{U}_x \times \mathcal{U}_y$, where restrictions are imposed.

Extending these considerations on several rules, we find these observations confirmed: Both figure 9 and figure 10 represent a fuzzy relation that is built of local statements. In the further, six rules of the form $[\tilde{A}_x \Rightarrow \tilde{B}_y]$ are interpreted as positive knowledge, whereas in the latter, six rules of the form $[\tilde{A}_x \Rightarrow \text{NOT} \tilde{B}_y]$ are interpreted as negative information. Consequently, the support distribution presented in figure 9 is the disjunction of six support distributions as the one in figure 6. Likewise, the possibility distributions depicted in figure 10 is built by conjunctively aggregating six possibility distributions like the one in figure 8.
Summarizing the above, we propose the following mechanism based on local possibilistic consequents to define a possibilistic rule base:

**How to specify a Gödel rule base:**

*In the possibilistic case, rule bases are sets of rules that exclude output states with respect to input states. In order to build such a rule base, for each linguistic term of the output variable, we specify all combinations of linguistic terms of the input variables that allow for excluding this very output.*

If we assume $k$ input variables and one output variable, the specification guidelines for Mamdani rule bases produce a rule base consisting of $O(n^k)$ rules, if there are $O(n)$ linguistic terms for each variable. In example 2, with five linguistic terms for each of the two input variables, we came up with $5^2 = 25$ rules, for instance.

At first glance, the possibilistic guideline seems to imply $O(n \cdot n^k)$ rules: For each of $O(n)$ linguistic terms of the output variable, we have to specify all those of the $O(n^k)$ possible combinations of the $k$ input variables that exclude the corresponding output term. Taking example 2 and considering only one output “it is $0_y$”, for instance, we already find $22 = O(5^2)$ combinations of inputs that fulfill the requirement to explicitly exclude the outcome “it is $0_y$”.

And still there are four more linguistic output terms left to deal with.

But fortunately we can combine several of those $O(n \cdot n^k)$ rules and come up with $O(n^k)$ rules like in the Mamdani case.

The following theorem shows the idea:

Theorem 1: In the possibilistic setting, the following two rule bases $R1$ and $R2$ are equivalent:

\[
R1 := \{ \text{IF } x \text{ is } A_1 \text{ AND } y \text{ is } B \text{ THEN } z \text{ is } C, \\
\quad \text{IF } x \text{ is } A_2 \text{ AND } y \text{ is } B \text{ THEN } z \text{ is } C \} 
\]

\[
R2 := \{ \text{IF } x \text{ is } A_1 \cup A_2 \text{ AND } y \text{ is } B \text{ THEN } z \text{ is } C \}
\]

Proof: Without loss of generality we assume for a fixed, but arbitrary triple $(u, v, w) \in U_x \times U_y \times U_z$

\[
\mu_{A_1}(u) \geq \mu_{A_2}(u). 
\]  

(11)

From this assumption we derive

\[
\Rightarrow \min \{ \mu_{A_1}(u), \mu_B(v) \} \geq \min \{ \mu_{A_2}(u), \mu_B(v) \}
\]

and this implies

\[
Gödel[\{A_1 \cup A_2 \land B \} \Rightarrow C] \leq \mu_{C}(z). 
\]

(12)

Applying the $\pi$-reasoning inference scheme as presented in section II, for the possibility distribution

\[
Gödel[\{A_1 \cup A_2 \} \land B \} \Rightarrow C] 
\]

representing rule base $R2$ it holds

\[
\forall u : \mu_{A_1 \lor A_2}(u) := \max \{ \mu_{A_1}(u), \mu_{A_2}(u) \}
\]

the possibility distribution $\mu_{A_1 \lor A_2}$

\[
\forall u : \mu_{A_1 \lor A_2}(u) := \begin{cases} 
1 & \text{if } \min \{ \mu_{A_1}(u), \mu_{A_2}(u) \} > 0, \\
\max \{ \mu_{A_1}(u), \mu_{A_2}(u) \} & \text{otherwise.}
\end{cases}
\]

This last expression represents the possibilistic rule base $R1$. 

Remark 1: Since both the linguistic conjunction AND and minimization $\min \{ \} \}$ are associative and commutative, the number and order of input variables does not matter in the proof of theorem 1, as long as just one variable is used for combination. This means that from any set of premise variables we can always arbitrarily choose one variable and apply the theorem in order to reduce the number of rules.

Example 3: Given this guideline, we can transform the Mamdani rule base from example 2 into a possibilistic one. On the fly we make use of theorem 1 in order end up with 25 instead of some 100 rules. Table II depicts the resulting rule base.

If now we apply the “critical” input $(x, y) = (2\tilde{x}_0, 3\tilde{y}_0)$ to this possibilistic rule base, we get the result depicted in figure 11. First of all, using this type of possibilistic rule base we get a reasonable result rather than the useless contradiction from figure 5 (right). Furthermore, this result is also better than the one from figure 5 (left), which is derived by positive knowledge and $\pi$-reasoning.

![Fig. 11. Applying input $(2\tilde{x}_0, 3\tilde{y}_0)$ to rule base of table II, $\pi$-reasoning (cf. figure 5)](image-url)
TABLE II
Gödel rule base from example 3

<table>
<thead>
<tr>
<th>IF x is AND y is</th>
<th>THEN z is</th>
</tr>
</thead>
<tbody>
<tr>
<td>10, y</td>
<td>30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>20, y</td>
<td>30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>30, y</td>
<td>30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>40, y</td>
<td>30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>50, y</td>
<td>30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>10, 2y</td>
<td>10, y ∪ 30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>20, 2y</td>
<td>10, y ∪ 30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>30, 2y</td>
<td>10, y ∪ 30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>40, 2y</td>
<td>10, y ∪ 30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
<tr>
<td>50, 2y</td>
<td>10, y ∪ 30, y ∪ 40, y ∪ 50, y NOT 30, y</td>
</tr>
</tbody>
</table>

we are able to make results more specific. From \( x \lor y \geq x \land y \) we derive

\[
Gödel[(\tilde{A}_x \lor \tilde{A}_2x) \Rightarrow \tilde{B}_y] \subseteq Gödel[(\tilde{A}_x \lor \tilde{A}_2x) \Rightarrow \tilde{B}_y].
\]

Applying drastic disjunction, overlapping premises tend to support each other, whereas non-overlapping premises are not affected. In general, this is reasonable from a semantic point of view: If we use a rule \([(\tilde{A}_x \lor \tilde{A}_2x) \Rightarrow \tilde{B}_y]\) instead of \([(\tilde{A}_1x \lor \tilde{A}_2x) \Rightarrow \tilde{B}_y]\), there are two possibilities:

1. \( \tilde{A}_1 \) and \( \tilde{A}_2 \) are non-neighboring linguistic terms and overlap each other, like \( \tilde{3} \) and \( \tilde{3} \) in example 2, for instance. In this case, it makes perfect sense to assume that the input \( x = 25 \) leads to the output \( "y \ is \ \tilde{B}" \), as well.
2. \( \tilde{A}_1 \) and \( \tilde{A}_2 \) are non-overlapping. In this situation, applying drastic disjunction instead of standard disjunction makes no difference at all.

V. Experiments

In the preceding section we proposed a simple guideline how to specify possibilistic rule bases. The proposal

is based on the idea that, as in the case of Mamdani inference, each rule should only provide local information. In this section we will present two examples, where using this guideline \( \pi \)-reasoning has clear advantages compared to \( \sigma \)-reasoning. As already pointed out in section III, due to inevitable accumulation of gradual inconsistencies and due to the way support distributions deal with inconsistent information, chaining of inferences leads to very broad and often useless results, when Mamdani inference is applied.

![Fig. 12. Subsequent results when chaining possibilistic rule base from example 4 (initial input \( x = 55 \)), cf. figure 2](image)

![Fig. 13. Final result when chaining possibilistic rule base from example 4 (initial input \( x = 50_x \)](image)

In example 1, section III, we examined those problems by chaining a fuzzy identity function using support distributions. Now we have a closer look at the very same example, this time applying possibilistic reasoning:

**Example 4:** Again we use a fuzzy rule base modeling the identity mapping \( y = f(x) \) with \( f(x) = x \). Using the same linguistic terms as used in example 1 and specified in figure 1, we now define an appropriate possibilistic rule base according to our guideline. For each linguistic term of the output variable \( y \) we specify all the input states that exclude this output term. Considering the first output term \( \tilde{1}_y \) and applying theorem 1, we get the following rule:

**If** \( x \) is \( \tilde{3}_x \lor \tilde{3}_x \lor \tilde{3}_x \lor \cdots \tilde{9}_x \lor \ \THEN \ \ y \ \ \text{is} \ \ \tilde{1}_y \).

Now we replace standard disjunction by drastic disjunction according to remark 2:

**If** \( x \) is \( \tilde{3}_x \lor \tilde{3}_x \lor \tilde{3}_x \lor \cdots \tilde{9}_x \lor \ \THEN \ \ y \ \ \text{is} \ \ \tilde{1}_y \).

Obviously, we can simplify this rule by using the premise "If \( x \) is NOT \( \tilde{1}_x \)" and consequently end up with the following rule base \( F_{\pi} \):

![image]
In this case, applying the initial input $\tilde{x} = 55$ we get a much more pleasing result than the one observed in $\sigma$-reasoning as depicted in figure 2. Furthermore, after the very first inference step, the result is stable. This is,

$$\tilde{y}_n = (F_\pi \circ (F_\pi \circ \ldots (F_\pi \circ \tilde{x}) \ldots)) = F_\pi \circ \tilde{x} = \tilde{y}_1$$

In figure 12, $\tilde{y}_1$ and $\tilde{y}_2$ are depicted. Like in the possibilistic example 3, the resulting constraint is quite reasonable and easy to defuzzify. Figure 13 presents the result of applying the fuzzy input $\tilde{x} = \tilde{y}_1$.

Figures 14 and 15 show the Mamdani rule base from example 1 and the Gödel rule base from example 4, respectively. Obviously, there are some differences in representing the identity mapping by means of the nine linguistic terms given, depending on the decision whether $\sigma$-reasoning or $\pi$-reasoning is applied.

On the other hand, now we pay a price for incomplete knowledge. The smaller triangles beneath the adequate major peak are caused by missing information. They represent values in output space, where consequents and corresponding premises are not very helpful, since they do not provide much support (Mamdani inference, $[\tilde{y}_0 \Rightarrow \tilde{y}_0]$) or impose severe restrictions (possibilistic case, $[\text{NOT} \tilde{y}_0 \Rightarrow \text{NOT} \tilde{y}_0]$). In order to emphasize this correlation, figure 16 replicates both the first inference result and the linguistic terms used to define the consequents of the corresponding rule base one upon the other. Obviously, the possibilistic result is not satisfying—namely: gradually incomplete—wherever the linguistic terms do not completely cover the output space.

Whereas causing some trouble in the possibilistic case, missing information does not play an observable role when $\sigma$-reasoning is applied, as to be seen in the first result $\tilde{y}_1$ of figure 2 on the upper left, for instance. Again, this comes as no surprise, since $\sigma$-reasoning simply does not care for missing information.

As stated several times before, we do not claim that either $\sigma$-reasoning or $\pi$-reasoning is the one and only mechanism to choose. As is clearly addressed in the discussion of the preceding paragraphs, both mechanisms have their advantages and drawbacks.

Nevertheless, since $\sigma$-reasoning in the disguise of Mamdani inference has gained so much attraction in practical applications, the main concern of this paper is to make $\pi$-reasoning a more attractive mechanism to apply. In order to do so we discuss a final, more complex experiment that favours the use of possibilistic reasoning.

**Example 5:** We model a simple expert system that is to guide dealing with company shares. From three inputs a
The personal demand PD tied by v e linguistic terms, as well: 

The three input variables are: (1) ES the evaluation of stock exchange, (2) EC the evaluation of company considered, and (3) FS the personal financial situation.

The knowledge base is built of two inference steps. First of all, considering the general situation at stock exchange ES and the special situation of the company considered EC, a personal demand PD to get company shares is derived. In the second stage, this subjective demand PD together with the financial situation FS are used to infer a final decision FD.

Both the stock exchange and the company are evaluated with respect to five linguistic terms, namely BAD, NEGATIVE, NORMAL, POSITIVE, and GOOD as depicted in figure 17.

The personal demand PD to get company shares is quantified by five linguistic terms, as well: --, --, +/-, +, and ++ (see figure 17) and increases both with ES and EC. The company evaluation EC is considered more important than the stock exchange evaluation ES. But when both ES and EC get too good at the same time, we do not want to get the company shares anymore, since everybody will. The corresponding Mamdani rule base to represent this relation is presented in table III.

The personal financial situation FS is quantified by means of five linguistic terms, namely BAD, MINUS, EVEN, PLUS, and GOOD (see figure 18). Both personal financial situation FS and personal demand PD have positive influence on the final decision NO, ??, YES, or YES!! This is, the Mamdani rule bases specified in table IV determines the final decision FD with respect to PD and FS.

The membership functions of the linguistic terms of the final decision FD are shown in picture 18.

In this example, the corresponding possibilistic rule bases are derived from their Mamdani counterparts listed above in the very same way that has been applied in example 3. Following the guideline for defining possibilistic rule bases, for any output term we determine all the input situations that exclude the output at hand. Applying both theorem 1 in order to combine linguistic terms of ES and remark 2, the \([ES \cap PD] \Rightarrow PD\) rule base is given in table V.

Table VI presents the possibilistic rule base \([FS \cap PD] \Rightarrow FD\) achieved in the very same way as the first rule base. Figure 19 shows typical results for both \(\sigma\)-reasoning (left) and \(\pi\)-reasoning (right) using the corresponding rule bases as introduced above. In each row, the same inputs have been applied to both mechanisms. Qualitatively, corresponding support and possibility distributions represent the same information. No matter whether we would use left or right hand sides of the figures — we would basically draw the same decisions.

On the other hand, possibilistic answers are more distinctive and thereafter more interesting. There is much more room for reasonable interpretation and discussion. Furthermore, if we need to use these results as input for subsequent rule base, the possibilistic answers seem much more appealing.

Compared to example 1, the results of this experiment are not too surprising: in both knowledge bases chaining takes place: fuzzy answers are used as inputs for subsequent fuzzy inference steps. And without any additional heuristics or intermediate defuzzification, Mamdani inference is
It was shown that a local positive piece of information turns out to be a major (global) restriction on the set of possible facts when interpreted in a possibilistic way. The positive information that the current temperature is 20 degrees, for example, means in a possibilistic setting that a temperature other than 20 degrees is impossible. Obviously the second statement is much stronger than the first one, i.e. the content of information is not the same.

VI. Conclusions

Mamdani inference has been very successful in control applications. This is not true for possibilistic reasoning. A closer look at both theories instantly reveals that they represent different kinds of knowledge: in case of Mamdani, facts are supported (positive knowledge) whereas possibility distributions exclude possible facts (negative knowledge). From this point of view it becomes clear that possibilistic and Mamdani rule bases have to be formulated in different ways. It was shown that a local positive piece of information turns out to be a major (global) restriction on the set of possible facts when interpreted in a possibilistic way. The positive information that the current temperature is 20 degrees, for example, means in a possibilistic setting that a temperature other than 20 degrees is impossible.
Another reason for the popularity of the Mamdani mechanism is the efficiency of its implementations. This does not hold for possibilistic inference. The hope for a fast local inference scheme is abandoned by the global effect of the Gödel inference relation. Fast implementations are based on approximations, i.e. [12]. A very promising method is introduced in [13]: the choice of an additive fuzzy system together with B-Splines as membership functions leads to a perfect duality of possibilistic and Mamdani inference. Computational costs become identically low, as well. Altogether it was shown that possibilistic inference based on Gödel inference relations is an equally important alternative to the classic Mamdani mechanism. Former problems were mainly a result of a general misunderstanding of how to formulate possibilistic information. As a solution a method for the transformation of positive into negative rules was introduced.

Acknowledgements

Important parts of this work were achieved during a stay of the first author at the International Computer Science Institute (ICSI) in Berkeley, California. The first author wants to thank the people at ICSI for providing excellent facilities and a great working atmosphere and the people of the BISC group – in particular Lotfi A. Zadeh – for contributing helpful feedback by means of questions, comments, and criticism.

References