A NEW METHOD OF PROVING THE EXISTENCE OF ANSWER SETS FOR DISJUNCTIVE LOGIC PROGRAMS: A METRIC FIXED POINT THEOREM FOR MULTI-VALUED MAPS

Mohamed A. Khamsi[1], Vladik Kreinovich[2], Driss Misane[3]
Departments of [1]Mathematics and [2]Computer Science
University of Texas at El Paso, El Paso, TX 79968
e-mail [1]mohamed@math.ep.utexas.edu [2]vladik@cs.ep.utexas.edu
[3]Universite Mohammed V, Faculte des Sciences
Departement de Mathematiques, Rabat, Morocco

Abstract. After the success of stable model semantics [GL88] and its generalization for programs with classical negations [GL90], the same ideas have been used in [GL91] to include disjunction into the resulting formalization of commonsense reasoning. The resulting semantics of disjunctive logic programs is relatively new, and therefore, few results are known. An additional problem with this semantics is as follows: since it incorporates a larger number of logical connectives, it is inevitably more complicated, and therefore, usual syntactic proofs are far less intuitive than the ones for stable models.

That these proofs become really complex, one can see from the fact that even for the simplest case of stratified programs, it is not so easy to prove that there always exists an answer set becomes really complicated. So, what is needed is a new methodology that would enable us to prove the existence of answer sets without going into syntactical details.

As a basis for this new methodology, we decided to use fixed point theorems. Traditionally, logic programming uses fixed point theorems a lot (see, e.g., [L87]), but these are fixed point theorems for monotonic mappings of ordered sets (to be more precise, Tarski's theorem on lattices). In disjunctive logic programming, an answer set can also be formulated as a fixed point, but this is a fixed point of a *multi-valued* mapping, and for such mappings neither Tarski's theorem, not other known result is applicable.

To overcome this difficulty, we decided to have a look at another case when Tarski's theorem was not directly applicable. Such a case was analyzed by M. Fitting in [F93] and [F93a]. Fitting showed that in some cases, we can use metric fixed point theorems to prove the existence of stable models. Fitting's theorem uses historically the first general metric fixed point theorem: so-called Banach's contraction principle [AK90], [K55]. This theorem is also not directly applicable to disjunctive logic programs, because it is about single-valued mappings, and we are dealing with multi-valued ones. However, Fitting theorem turned out to be a prefect starting point for us.

In this paper, we describe a generalization of the contraction principle to multi-valued mappings, and show that the resulting generalization can be used to produce a simple proof that every stratified disjunctive logic program has an answer set. This result is easily generalizable to locally stratified disjunctive logic programs [P88].

1. BASIC DEFINITIONS

General remark. Since the main purpose of this paper is to promote a new methodology, we have tried to make it as understandable as possible. For that reasons, we are including all the definitions; readers who are already familiar with these definitions, can skip the corresponding subsection.

1.1. What is a disjunctive logic program and what is an answer set [GL91]

Remark. The definition of an answer set follows the tradition of a stable model semantics in that it is done in two steps:

- first, if we have variables in the original logic program, we substitute all possible terms instead of them; as a result, we get a new program that contains only ground instances of all the rules;
- after that, we apply a special procedure to this new logic program to check whether a guess is an answer set (or, in case of a stable model semantics, a stable model).

Because of that, when defining answer set, we can safely assume that our program already has no atoms and contains only ground instances of the rules. Thus, we arrive at the following definition.

Definition 1 [GL91]. Assume that a set $\mathcal{A}$ is given. Its elements will be called *atoms*. By a *literal* we mean either an atom, or an expression of the type $\neg p$, where p is an atom. This symbol $\neg$ will be called *classical negation*. The set of all literals is denoted by *Lit*. An *extended rule* (or *rule* for short) is an expression of the form

$$L_1 \vee L_2 \vee ... \vee L_k \longleftarrow L_{k+1} \wedge L_{k+2} \wedge ... \wedge L_{k+m} \wedge \mathit{not}\ L_{k+m+1} \wedge ... \wedge \mathit{not}\ L_{k+m+n},$$

where n, m, k are non-negative integers. We will say that literals $L_1, ..., L_k$ are *in the head*, and literals $L_{k+1}, ..., L_{k+m+n}$ are *in the body* of this rule. An *extended disjunctive logic program* Π is a set of extended rules.

Remarks.
1. It is usually assumed that the set *Lit* coincides with all the literals that can be formulated in the language of the original logic program (with variables).

2. A rule with $k = 1$ and $m = n = 0$ is called a *fact.*

Definition 2 (of an answer set) [GL91]
1. Let Π be an extended disjunctive program which doesn't contain *not.* An *answer set* of Π is a minimal set $S \subseteq \mathit{Lit}$ such that
 (i) for each rule $L_1 \vee L_2 \vee ... \vee L_k \longleftarrow L_{k+1} \wedge L_{k+2} \wedge ... \wedge L_{k+m}$ from Π, if $L_{k+1}, .., L_{k+m} \in S$, then, for some $i \in \{1, .., k\}$, $L_i \in S$ (sets that satisfy (i) are called *closed under the rules* of Π).
 (ii) if S contains a pair of complimentary literals (a and $\neg a$), then $S = \mathit{Lit}$.

2. Now, let Π be an arbitrary extended disjunctive logic program. For any set $S \subseteq \mathit{Lit}$, let Π^S denote the extended disjunctive program obtained from Π by deleting
 (i) each rule that has a formula *not* L in its body with $L \in S$, and

(ii) all formulas of the form *not* L in the bodies of the remaining rules.

Clearly Π^S doesn't contain *not*, so for this program, we can use part 1 of this Definition to define the set of all its answer sets. We will denote this set by $\alpha\left(\Pi^S\right)$. If $S \in \alpha\left(\Pi^S\right)$, then we say that S is an *answer set* for Π.

1.2. What is a stratified disjunctive logic program

Definition 3. Let Π be an extended logic program. We will say that Π is *stratified* if for some integer $\alpha > 1$,

$$Lit = \bigcup_{1 \le i < \alpha} Lit_i,$$

where Lit_i are disjoint sets called *strata*, so that for every rule

$$L_1 \vee L_2 \vee ... \vee L_k \longleftarrow L_{k+1} \wedge L_{k+2} \wedge ... \wedge L_{k+m} \wedge not\ L_{k+m+1} \wedge ... \wedge not\ L_{k+m+n},$$

in Π, the following two statements are true:
(i) if L_i is in the head, L_j is in the body, $L_i \in Lit_a$, and $L_j \in Lit_b$, then $a \ge b$;
(ii) if L_i is in the head, *not* L_j is in the body, $L_i \in Lit_a$, and $L_j \in Lit_b$, then $a > b$.

A decomposition $\{Lit_i\}$ of Π satisfying the above conditions is called a *stratification* of Π.

Remark. Following [P88], we can generalize this definition to the case when α is an arbitrary countable ordinal. Such a program is called *locally stratified*, and the corresponding decomposition is called a *local stratification*. In particular, we can take $\alpha = \omega$. We will call locally stratified programs with $\alpha = \omega$, *countably stratified*, and the corresponding decomposition

$$Lit = \bigcup_{i=1}^{\infty} Lit_i$$

will be called a *countable stratification*.

1.3. Metric spaces: main definitions, contraction theorem [K55]

Definition 4. Let M be an arbitrary set. A *metric* or *distance* on M is a mapping $d : M \times M \to \mathbf{R}_+$ (where $\mathbf{R}_+$ denotes the set of all non-negative real numbers) such that
(i) $d(x, y) = 0$ if and only if $x = y$;
(ii) $d(x, y) = d(y, x)$;
(iii) $d(x, y) \le d(x, z) + d(z, y)$.
The distance d will be said to be *ultrametric* if it satisfies the additional inequality
(iv) $d(x, y) \le \max(d(x, z), d(z, y))$.
A pair (M, d) of a set M and a distance d defined on M is called a *metric space*.

To describe contraction principle, we need a few extra definitions.

Definition 5. Let M be a metric space.

(i) A sequence $\{x_n\}$ of elements of M is said to be *convergent* to $x \in M$ if

$$\forall \varepsilon > 0 \exists n_0 \forall n (n \geq n_0 \implies d(x_n, x) < \varepsilon).$$

(ii) A sequence $\{x_n\}$ is said to be *Cauchy* if

$$\forall \varepsilon > 0 \exists n_0 \forall n, m (n \geq n_0 \;\&\; m \geq n_0 \implies d(x_n, x_m) < \varepsilon).$$

(iii) A metric space M is said to be *complete* if all Cauchy sequences in M are convergent.

(iv) A mapping $T : M \to M$ is called a *contraction* if there exists a real number $k < 1$ such that $d(T(x), T(y)) \leq k d(x, y)$ for every $x, y \in M$.

(v) A point $x \in M$ is called a *fixed point* of a mapping T if $T(x) = x$.

BANACH'S CONTRACTION PRINCIPLE [K55], [AK90]. *If M is a complete metric space, then every contraction on M has a unique fixed point.*

Remarks.

1. The proof of this Theorem is rather constructive: we start with an arbitrary point $x_0 \in M$, and define $x_{n+1} = T(x_n)$ for $n = 0, 1, 2, \ldots.$ Then, from the contraction property, we conclude that $d(x_{n+1}, x_n) \leq k d(x_n, x_{n-1})$; therefore, $d(x_{n+1}, x_n) \leq k^n d(x_0, T x_0)$. From this, one can conclude that the sequence $\{x_n\}$ is Cauchy, and that its limit x is a fixed point.

2. This Theorem was used by M. Fitting to prove the existence of a stable model ([F93], [F93a]). Fitting does not apply his result directly to stratified programs; however, he considers the general case when we have a *level mapping* l, i.e., in our denotations, a mapping from *Lit* to the set N of all positive integers. This mapping means that we actually have a decomposition of *Lit* into disjoint sets $S_i = \{L \in Lit \,|\, l(L) = i\}$. Vice versa, if we have a decomposition into countably many sets S_i, we can define a level mapping l as follows: $l(L)$ is the number i of the only set S_i to which L belongs. Stratification is a particular case of such a partition. For stratification, Fitting's definition reduces to the following:

Definition 6. Let Π be a countably stratified extended disjunctive program. We will then define a *Fitting metric* d on the set 2^{Lit} of all subsets of *Lit* as follows:

- if $A = B$, then $d(A, B) = 0$;
- if $A \neq B$, then $d(A, B) = 2^{-(m-1)}$, where m is the smallest integer for which $A \cap Lit_m \neq B \cap Lit_m$.

Remark. It is well known that a stable model can be defined as a fixed point of some mapping $T : 2^{Lit} \to 2^{Lit}$. Fitting actually proved that the set $M = 2^{Lit}$ with this metric d is complete, and that the corresponding T is a contraction with $k = 1/2$. Thus, he proved the existence of a fixed point (i.e., the existence of a stable set).

PROPOSITION 1 [F93], [F93a]. *The set M with Fitting's metric d is a complete metric space.*

Remark. One can easily prove that Fitting's metric is an ultrametric.

2. HOW TO APPLY METRIC FIXED POINTS TO DISJUNCTIVE LOGIC PROGRAMS

The idea. The stable set S is defined as a fixed point of some mapping $T : M \to M$, i.e., as a set for which $S = T(S)$. The answer set was defined as a set S for which

$$S \in \alpha(\Pi^S) \tag{1}$$

Here, the mapping $S \to \alpha(\Pi^S)$ assigns to every element $S \in M$, a *set* of elements of M. Such mappings are called *multi-valued*. We can thus view formula (1) as a (natural) generalization of the notion of a fixed point to multi-valued mappings. Let's give formal definitions.

Definition 7. By a *multi-valued* mapping of a set M into itself, we mean a mapping $T : M \to 2^M$. We say that an element $x \in M$ is a *fixed point* of a multi-valued mapping T if $x \in T(x)$.

COROLLARY. A set S is an answer set of an extended disjunctive logic program Π iff it is a fixed point of the multi-valued mapping $S \to \alpha(\Pi^S)$.

Remark. Since this T is a multi-valued mapping, it cannot be a contraction, and therefore, we cannot directly apply Banach's contraction principle to it. Let us find what we *can* prove in this case. It turns out that the same ideas that Fitting used to prove that his T was contraction, lead to the following result:

PROPOSITION 2. *Let Π be a countably stratified extended disjunctive logic program. Let $S_1, S_2 \in M = 2^{Lit}$. Then*

$$\forall A \in \alpha\Big(\Pi^{S_1}\Big) \exists B \in \alpha\Big(\Pi^{S_2}\Big) \text{ such that } d(A,B) \le \frac{1}{2} d(S_1, S_2).$$

Proof. The main idea of this proof is the same as in the Fitting's proof that his T is a contraction mapping: namely, if $d(A,B) = 2^{-n}$, this means that the restriction of A to a strata Lit_i ($1 \le i \le n$) coincides with the restriction of B to the same strata. Since we are dealing with a stratified program, for an arbitrary literal $L \in Lit_{n+1}$, whether we use A or B as an initial guess, all the rules that have L in the head will have exactly the same parts of their bodies deleted. Therefore, exactly the same rules with no negation as failure will determine the validity of L. Hence, the resulting sets of answer sets will agree on all $L \in L_{n+1}$. In other words, if $S_A \in \alpha(\Pi^A)$, then there exists $S_B \in \alpha(\Pi^B)$ for which $S_A \cap L_{n+1} = S_B \cap L_{n+1}$. Hence, $d(S_A, S_B) \le 2^{-(n+1)}$. Q.E.D.

Remark. Our idea is to prove the desired result (that every stratified disjunctive logic program has an answer set) by proving that every mapping with this "contraction" property has a fixed point. To the best of our knowledge, none of the existing fixed point theorems (see, e.g., [AK90]) is directly applicable to this case. So, we must prove a new theorem to cover it. Here is the theorem that we proved (and that covers the case of disjunctive logic programs):

Definition 8. Let's say that a multi-valued mapping $T : M \to 2^M$ is a *contraction* if there exists a real number $k < 1$ such that for every $x \in M$, for every $y \in M$, and for all $a \in T(x)$, there exists $b \in T(y)$ such that $d(a, b) \le kd(x, y)$.

FIXED POINT THEOREM FOR MULTIPLE-VALUED CONTRACTIONS. *Assume that M is a complete metric space, T is a multi-valued contraction on M such that $T(x)$ is not empty for some $x \in M$ (i.e., T is not identically empty), and for every $x \in M$, the set $T(x)$ is closed. Then, T has a fixed point.*

Remark. An identically empty mapping $T : M \to 2^M$, for which $T(x) = \phi$ for every x, evidently, cannot have fixed points (because $x \notin \phi = T(x)$), so the additional condition that T is not identically empty is necessary.

Proof. Since T is not identically empty, there exists an element $x_0 \in M$ such that $T(x_0) \ne \phi$. Let $x_1 \in T(x_0)$. Since T is a contraction (in the sense of Definition 8), we insure the existence of $x_2 \in T(x_1)$ such that $d(x_1, x_2) \le kd(x_0, x_1)$. We can apply the same argument again, and thus step-by-step we will construct a sequence x_n such that for every $n \ge 0$, we have $x_{n+1} \in T(x_n)$ and $d(x_{n+1}, x_{n+2}) \le kd(x_n, x_{n+1})$.

Using the triangle inequality, one can prove that

$$d(x_n, x_{n+m}) \le \sum_{i=0}^{m-1} d(x_{n+i}, x_{n+i+1}) \le \sum_{i=0}^{m-1} k^{n+i} d(x_0, x_1) \le \frac{k^n}{1-k} d(x_0, x_1).$$

This implies that $\{x_n\}$ is a Cauchy sequence in M. Since M is complete, there exists a point x_ω which is the limit of x_n. From the above inequality, we conclude in the limit $m \to \infty$ that

$$d(x_n, x_\omega) \le \frac{k^n}{1-k} d(x_0, x_1). \tag{2}$$

Let us show that x_ω is a fixed point of T.

Since T is a contraction, for every n, there exists $y_n \in T(x_\omega)$ such that $d(x_{n+1}, y_n) \le kd(x_n, x_\omega)$. Therefore, $d(x_\omega, y_n) \le d(x_\omega, x_{n+1}) + d(x_{n+1}, y_n) \le d(x_\omega, x_{n+1}) + kd(x_n, x_\omega)$. Because of (2), we now have

$$d(x_\omega, y_n) \le \frac{k^{n+1}}{1-k} d(x_0, x_1) + k \frac{k^n}{1-k} d(x_0, x_1) = \frac{2k^{n+1}}{1-k} d(x_0, x_1).$$

In the limit $n \to \infty$, $d(x_\omega, y_n) \to 0$, hence $y_n \to x_\omega$.

But $y_n \in T(x_\omega)$, and $T(x)$ is closed for every x. Therefore, the limit x_ω of the sequence y_n also belongs to $T(x_\omega)$. So, $x_\omega \in T(x_\omega)$, and x_ω is the desired fixed point. Q.E.D.

To apply this general theorem to disjunctive logic programs, we thus need to prove two lemmas:

LEMMA 1. *For a countably stratified disjunctive logic program* Π*, and for every set* $S \in 2^{Lit}$*, the set* $\alpha(\Pi^S)$ *is closed with respect to Fitting's metric.*

Proof. Let $S \in 2^{Lit}$ and $\{A_n\}$ be a sequence of elements in $\alpha(\Pi^S)$ which converges to A. Let us prove that $A \in \alpha(\Pi^S)$.

To prove that, we need to prove the two things:
- that A is closed under the rules of Π^S, and
- that A is minimal, i.e., that no $B \subset A$, $B \neq A$, is closed under these rules.

1. Let us first prove that A is closed under the rules (i.e., if all the literals from the body of a rule belong to A, then one of the literals from the head of this rule also belongs to A). Let R be a rule in Π^S such that every literal from its body belongs to A. Let's prove that one of the literals from its head also belongs to A.

Indeed, every literal from R (i.e., from its body and from its head) belongs to some stratum Lit_i (for some integer i). Let us denote the maximum of these integers i by s (s for *s*tratum). Then, all literals L from the body of A belong to the union of the first s strata:

$$L \in \bigcup_{i=1}^{s} Lit_i.$$

Since A_n converges to A, there exists an n_0 such that for all $n \geq n_0$, we have $d(A_n, A) \leq 2^{-(s+1)}$. In particular, for $n = n_0$, $d(A_{n_0}, A) \leq 2^{-(s+1)}$. By definition of Fitting's metric d, this means that $A \cap Lit_i = A_{n_0} \cap Lit_i$ for all $i \leq s$. We have chosen s in such a way that all literals from R belong to one of the strata Lit_i, $i \leq s$. According to our choice of R, every literal L from the body of R belong to A. Therefore, it belongs to $A \cap Lit_i$ for some $i \leq s$. Since $A \cap Lit_i = A_{n_0} \cap Lit_i$, this literal L thus belongs to $A_{n_0} \cap Lit_i$ and hence, to A_{n_0}.

So, all literals from the body of R belong to A_{n_0}. But $A_{n_0} \in \alpha(\Pi^S)$ and therefore, A_{n_0} is closed under the rules. Hence, one of the literals from the head of R also belongs to A_{n_0}. Because of our choice of s, this literal L belongs to Lit_i for some $i \leq s$. Therefore, $L \in A_{n_0} \cap Lit_i = A \cap Lit_i$, and hence $L \in A$. So, A is closed.

2. Let us now show that A is minimal.

We will prove it by reduction to a contradiction. Assume that A is not minimal. This means that there exists a set $B \subset A$, $B \neq A$, that is closed under the rules of Π^S.

The fact that $B \neq A$ means that some literal L from A is not in B. Let us denote the number of the stratum to which this literal L belongs, by s (i.e., $L \in Lit_s$). Then, $B \cap Lit_s \neq A \cap Lit_s$. Since $A_n \to A$, there exists an n_0 such that for all $n \geq n_0$, we have $d(A_n, A) \leq 2^{-(s+1)}$. In particular, for $n - n_0$, we have $d(A_{n_0}, A) \leq 2^{-(s+1)}$. Similarly to Part 1 of this proof, this means that $A_{n_0} \cap Lit_i = A \cap Lit_i$ for all $i \leq s$.

Let us now define a new set

$$B^* = (B \bigcap \bigcup_{i=1}^{s} Lit_i) \ \bigcup \ (A_{n_0} \bigcap \bigcup_{i=s+1}^{\infty} Lit_i)$$

It is easy to see that B^* is closed under the rules of Π^S.

For $i \leq s$, from $B \subset A$, we conclude that $B^* \cap Lit_i \subseteq A \cap Lit_i = A_{n_0} \cap Lit_i$ and hence, $B^* \cap Lit_i \subseteq A_{n_0} \cap Lit_i$. For $i > s$, a similar inclusion $B^* \cap Lit_i \subseteq A_{n_0} \cap Lit_i$ follows directly from the definition of B^*. So, for every stratum, its intersection with B^* is a subset of its intersection with A_{n_0}. Therefore, $B^* \subseteq A_{n_0}$. On the other hand, for $i = s$ (because of our choice of s) $B^* \cap Lit_s = B \cap Lit_s \neq A \cap Lit_s = A_{n_0} \cap Lit_s$. Therefore, $B^* \neq A_{n_0}$. So, inside A_{n_0}, there is another set that is also closed under the rules of Π^S. So, A_{n_0} is not minimal.

But $A_{n_0} \in \alpha(\Pi^S)$ and therefore, A_{n_0} is minimal. This contradiction shows that our assumption (that A is not minimal) is false. Hence, A is minimal.

3. So, A is closed under the rules, and A is minimal. Hence $A \in \alpha(\Pi^S)$. Q.E.D.

LEMMA 2. *For a countably stratified disjunctive logic program Π, there exists a set $S \subseteq Lit$ such that $\alpha(\Pi^S)$ is not empty.*

Remark. To prove this Lemma, we will prove the following more general result.

LEMMA 3. *For a countably stratified disjunctive logic program Π, for every set $S \subseteq Lit$, $\alpha(\Pi^S)$ is not empty.*

Proof. For an arbitrary set S, the set Lit is closed under the rules of Π^S. One can easily prove that there exists a minimal set with this property. Therefore, $\alpha(\Pi^S)$ is not empty. Q.E.D.

Lemmas 1 and 2 enable us to prove the main result:

MAIN RESULT. *Let Π be a countably stratified extended disjunctive logic program. Then Π has an answer set.*

Since stratified programs are a particular case of countably stratified programs, we have the following Corollary:

COROLLARY. *Let Π be a stratified extended disjunctive logic program. Then Π has an answer set.*

Remarks.

1. For locally stratified disjunctive logic programs (with an arbitrary countable ordinal α), we can apply a similar proof, with the only difference that instead of Fitting's metric d, we will have to consider an "ultrametric" whose values are not real numbers, but elements of the ordered set $\{2^{-\beta}\}$, where β are all ordinals $< \alpha$, and the order is defined as follows: $2^{-\beta} < 2^{-\gamma}$ iff $\beta > \gamma$. This result is somewhat technical, so we describe it in all necessary detail in the Appendix.

2. Fitting in [F93], [F93a] provides an example when a metric fixed point theorems proves the existence of a stable set for a non-stratified program. Namely, he considers a program that describes wining positions in a positional game, in which players make moves in turn

and there are no draws. If a position is winning for one of the players, then no further moves are possible. This program consists of the rules and the facts. Rules are of the following type:

$$win(X) \leftarrow move(X, Y), not\ win(Y),$$

and the facts (of the type $move(X, Y)$) describe possible moves. Predicate $win(X)$ means that X is a winning position, i.e., if a player is in a position X, then there exists a strategy that enables him to win (no matter what the actions of the opposite player are).

The (informal) meaning of the above rule is as follows: if a player is in a position X, and he can move into another position Y that is not winning (i.e., losing) for the opposite player, then he wins. If he cannot make such a move, this means that wherever he moves to, the resulting position is winning for the second player. So, if no such move exists, then X is a losing position for X.

Let's modify this example into an example where a metric fixed point theorem helps to find a stable model for a non-stratified disjunctive logic program. We will consider the same game, but this time, we will assume that the players are still training (it is not yet a championship). So, if a person is about to win, then instead of going all the way to his victory, he can stop the game and teach another player (i.e., explain how he could win). A program that describes this situations contains the same facts $win(X, Y)$, but slightly different rules:

$$can\,win(X) \leftarrow move(X, Y), not\ can\,win(Y)$$

$$win(X) \vee teach(X) \leftarrow can\,win(X).$$

Rules of the first type describe when a player can win. Rules of the second type tell that if a player can win, then he will either win, or teach. This program is non-stratified. However, for this program, the mapping $S \rightarrow \alpha(\Pi^S)$ is a contraction (this is proved just like in [F93] and [F93a]) and therefore, our fixed point theorem proves the existence of its fixed point (i.e., of a stable model of Π).

Acknowledgments. This work was partially supported by an NSF grant No. CDA-9015006. The authors are thankful to M. Gelfond and C. Baral for the formulation of the problem, and for valuable discussions of its logic programming aspects, to I. Rival for discussing the fixed point part, and to M. Fitting for encouragement and important references.

REFERENCES

[AK90] A. G. Aksoy and M. A. Khamsi, *Nonstandard methods in fixed point theory*, Springer-Verlag, N.Y., 1990.

[F93] M. Fitting, *Metric methods: three examples*, Technical Report, City University of New York, 1993.

[F93a] M. Fitting, *Metric methods, three examples and a theorem*, Journal of Logic Programming, 1993 (to appear).

[GL88] M. Gelfond and V. Lifschitz, *The stable model semantics for logic programming*, In: R. Kowalski and K. Bowen (eds.), *Logic programming: Proceedings of the 5th Intl. Conference and Symposium*, 1988, pp. 1070–1080.

[GL90] M. Gelfond and V. Lifschitz, *Logic programs with classical negation*, In: D. Warren and P. Szeredi, *Logic Programming: Proceedings of the 7th International Conference*, 1990, pp. 579–597.

[GL91] M. Gelfond and V. Lifschitz, *Classical negation in logic programs and disjunctive databases*, New Generation Computing, 1991, Vol. 9, pp. 365–385.

[K55] J. Kelley, *General topology*, Van Nostrand, Princeton, NJ, 1955.

[L87] J. W. Lloyd, *Foundations of logic programming*, Springer-Verlag, N.Y., 1987.

[P88] T. C. Przymusinski, *On the declarative semantics of deductive databases and logic programs*, In: J. Minker (ed.), *Foundations of deductive databases and logic programming*, Morgan Kaufmann, Los Altos, CA, 1988.

APPENDIX: GENERALIZED METRICS, FIXED POINTS, AND THE EXISTENCE OF ANSWER SETS FOR LOCALLY STRATIFIED DISJUNCTIVE LOGIC PROGRAMS

A1. MOTIVATIONS FOR THE FOLLOWING DEFINITIONS

A1.1. Why it is necessary to generalize the notion of a metric space

Since metric fixed points can be applied so naturally to stratified logic programs, it makes sense to try to apply similar techniques to *locally* stratified logic programs. The first thing we need is to generalize Fitting's metric to such programs. Fitting's distance is defined as follows: $d(A,B)=2^{-(m-1)}$, where m is the smallest integer for which $A \cap Lit_m \neq B \cap Lit_m$. For locally stratified programs, we have strata Lit_m indiced by ordinal numbers m. So, in addition to integer m, we have values m that are infinite ordinals (not integers). If the first strata Lit_m for which $A \cap Lit_m \neq B \cap Lit_m$ is an infinite ordinal, then, if we literally follow Fitting's definition, we will define $d(A,B)$ as $2^{-\alpha}$ for some infinite ordinal α. This is just an expression, because the operation 2^{-x} is not defined for infinite ordinals x. So, to apply Fitting's idea, we must make sense out of this expression.

Since this $2^{-\alpha}$ is the distance, maybe we can somehow interpret is as a non-negative real number? Alas, such an interpretation is impossible. To be more precise, we can interpret it as whatever we want, but if we do interpret is as a non-negative real number, the above-given proof will be no longer applicable. Indeed, one of the properties of 2^{-m} that we used in our proofs was that if $m<n$ then $2^{-m}>2^{-n}$ (i.e., that a function $x \to 2^{-x}$ is strictly decreasing, or, using the term used in logic programming community, *antitonic*). Since we want to use a similar proof, we would like to keep this property. If α is an infinite ordinal, then $\alpha>n$ for every integer n, and therefore, since we want monotonicity, we would have $2^{-\alpha}<2^{-n}$ for all n. So, if $2^{-\alpha}$ is a real number, then we can conclude that $2^{-\alpha} \leq 0$. On the other hand, $2^{-\alpha}$ is a non-negative real number, so the only possibility is that $2^{-\alpha} = 0$. But according to the definition of a metric, $d(A, B) = 0$ if and only if $A = B$. So, if we take two sets A and B that coincide for all the strata Lit_m for finite m and differ for some strata Lit_α, then on one hand, these two sets are different: $A \neq B$, and on the other hand, $d(A, B) = 2^{-\alpha} = 0$ and so, $A = B$. This contradiction shows that we cannot use real numbers.

Instead, we need a "metric" with the values in some other set that will include not only real numbers, but also some non-negative elements that are smaller than 2^{-n} for all n. Such sets are known (in particular, in the so-called *non-standard analysis*), and the corresponding elements (that are smaller than any positive fraction but still greater than 0) are called *infinitesimal* (see, e.g., [AK90]).

A1.2. What properties of the generalized metric do we need? A general description

In order to handle the locally stratified case (and hopefully, for some future applications as well), we must describe a (generalized) metric on a set M as a mapping from the set of pairs $M \times M$ to some set V (V for *v*alues). We want to be able to apply standard results about metrics to generalized metrics as well. Therefore, we want a generalized metric to satisfy properties (i)–(iii) from Definition 4.

By simply looking at these properties, we can immediately find out that in order to be able to formulate these properties, this set V must be equipped with the following additional structure:
- V must contain an element called 0;
- an ordering relation $<$ must be defined on V;
- a binary operation $+$ must be defined on V.

What properties do we want these structures to satisfy?

Let us start with $+$.
- In the usual metric spaces, for $y = z$, triangle inequality is trivially true, because it turns into $d(x, y) \leq d(x, y) + d(y, y) = d(x, y) + 0$, and for usual addition, $v + 0 = a$ for all v. So, we would like to retain this property.
- When we formulate triangle inequality for the usual metric, we do not care whether we write it in a way is written in Definition 4 ($d(x, y) \leq d(x, z) + d(z, y)$), or in the form $d(x, y) \leq d(z, y) + d(x, z)$, because for usual addition, $u + v = v + u$ for all u, v (i.e., $+$ is a *commutative* operation). In our generalization, we want to be able to simply apply standard metric results, without thinking about what order is appropriate. So, we want our operation $+$ to be also commutative.
- For usual metric spaces, from triangle inequality, we can easily deduce a more general inequality, e.g., $d(x, y) \leq d(x, z) + d(z, t) + d(t, y)$. Here, we do not to worry in what order to apply add (i.e., whether we should understand the right-hand side of this inequality as $(d(x, z) + d(z, t)) + d(t, y)$ or $d(x, z) + (d(z, t) + d(t, y))$ because for normal addition, it does not matter: addition is *associative* ($(u+v)+w = u+(v+w)$ for all u, v, and w). We do not want to worry about the order for generalized metric, so we want the operation $+$ on V to be associative as well.

A set V with an associative binary operation $+$ is called a *semigroup*. If $+$ is also commutative, this semigroup is called *commutative*, or *Abelian*. If an Abelian semigroup has an element 0 with the property that $0 + v = v + 0 = v$ for all v, then this element is called a *zero*, and V is called an *Abelian semigroup with a 0*.

Now, about the ordering relation $<$:

- For usual metric, triangle inequality is trivially true for $x = y$, because then, it turns into $0 = d(x, x) \leq d(x, y) + d(y, x)$, and $0 \leq v$ for any element $v \in \mathbf{R}_+$. We would like to retain this property for V, i.e., we would like to have $v \geq 0$ for every $v \in V$.
- Arguments about metric often use the fact that if $v_1 \leq v_2$ and $\tilde{v}_1 \leq \tilde{v}_2$, then $v_1 + \tilde{v}_1 \leq v_2 + \tilde{v}_2$. So, we would also like to retain this property.

A semigroup with an order that agrees with $+$ (in this sense) is called an *ordered semigroup*.

A1.3. The set of values that we will use

We want to describe the case when "distances" are described by "numbers" $2^{-\alpha}$ for different ordinals α. Here, the ordering is natural: $2^{-\alpha} < 2^{-\beta}$ iff $\alpha > \beta$.

Fitting's metric is actually an ultrametric, i.e., triangle inequality is true for max instead of $+$. Therefore, as a desired semigroup operation, we can take max.

A1.4. How to define convergence for generalized metric spaces?

For usual metric, convergence can be defined in two different ways: First, we can define it as in Definition 5: a sequence $\{x_n\}$ *converges* to $x \in M$ if a sequence d_n, defined as $d_n = d(x_n, x)$, tends to 0 in the sense that $\forall \varepsilon > 0 \exists n_0 \forall n (n \geq n_0 \implies d_n < \varepsilon)$. This is a traditional definition (that uses a so-called "$\varepsilon - \delta -$language").

This definition has a perfect computational sense. Suppose that we are interested in some number x (e.g., x is an (unknown) solution to a given equation). In many cases, we cannot immediately produce an exact expression for x. Instead, we have an iterative procedure that computes x with better and better accuracy. If by x_n, we denote the result of $n-$th iteration, then $x_n \to x$. This means that if we want to compute x_n with some accuracy $\varepsilon > 0$, then for some n_0, our iterative procedure enables us to do that.

From computational viewpoint, however, convergence has no practical value unless we know n_0. Indeed, if we do not know n_0, then no matter where we stop, there is not guarantee that we actually computed x with a desired accuracy. So, to make convergence meaningful, we must not only postulate that for every ε, there exists an n_0, but also provide an algorithm, that describes an accuracy v_n of step n. In other words, we must have a decreasing sequence $v_n \downarrow 0$ such that $d_n \leq v_n$ for all n.

How can we notice that iterations are converging? A usual way of seeing that is by noticing that the results obtained on several iterations are very close to each other, i.e., that the distance $d(x_n, x_m)$ between these results becomes small when m and n increases.

How small can it become? If $d(x_n, x) \leq v_n$, and $d(x_m, x) \leq v_m$, then because of triangle inequality, $d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) \leq v_n + v_m$. Vice versa, if the inequality

$d(x_n, x_m) \le v_n + v_m$ is true for all n and m, then, when we go to the limit $m \to \infty$, we can conclude that $d(x_n, x) \le v_n$. The usual notion of a Cauchy sequence describes when a sequence converges. The above inequality tells when a sequence converges with a given order of convergence. So, this inequality is a reasonable (algorithms-inspired) analogue of the notion of a Cauchy sequence for convergence with a given order.

We consider the case when a rapidly decreasing sequence (like 2^{-n}) not necessarily converges to 0. To get 0, it is thus not sufficient to get a sequence of iterations: we may need to use additional steps after infinitely many iterations. In mathematical terms, we must use *transfinite induction*.

Now, we are ready for the formal definitions.

A2. GENERALIZED METRIC SPACES: DEFINITIONS AND THE MAIN RESULT

A2.1. Definitions for the general case

Definition A1. By a *semigroup*, we mean a set V with an associative binary operation $+ : V \times V \to V$. If $+$ is also commutative, then a semigroup is called *commutative*, or *Abelian*. A semigroup is called a *semigroup with 0*, if there exists an element $0 \in V$ such that $0 + u = u$ for all $u \in V$.

Definition A2. By an *ordered semigroup with 0*, we mean a semigroup with 0, on which there is an ordering $<$ such that $0 \le v$ for all $v \in V$, and if $v_1 \le v_2$ and $\tilde{v}_1 \le \tilde{v}_2$, then $v_1 + \tilde{v}_1 \le v_2 + \tilde{v}_2$.

Definition A3. Let V be an ordered Abelian semigroup with 0, and let M be an arbitrary set. A *generalized metric* (or *generalized distance*) on M is a mapping $d : M \times M \to V$ such that
(i) $d(x, y) = 0$ if and only if $x = y$;
(ii) $d(x, y) = d(y, x)$;
(iii) $d(x, y) \le d(x, z) + d(z, y)$.
A pair (M, d) of a set M and a distance d defined on M is called a *generalized metric space*.

A2.2. Our set of values

Comment. In the following text, we will use the following set of values:

Definition A4. Let us denote by V the set of all expressions of the type 0 or $2^{-\alpha}$, where α is a countable ordinal (i.e., $\alpha < \omega_1$, where ω_1 denotes the first non-countable ordinal). An order on V is defined as follows: $0 \le v$ for every $v \in V$, and $2^{-\alpha} < 2^{-\beta}$ iff $\alpha > \beta$. As a semigroup operation $+$, we will use the maximum $\max(u, v)$.

A2.3. Convergence

Definition A5. Assume that α is either a countable ordinal or ω_1, and $\mathbf{v} = \{v_\beta\}_{\beta<\alpha}$ is a decreasing family of elements of V. Let M be a generalized metric space, and let $\{x_\beta\}_{\beta<\alpha}$ be a family of elements of M.

(i) A family $\{x_\beta\}$ is said to $\mathbf{v}-$*cluster* to $x \in M$ if

$$\forall\beta(\beta < \alpha \implies d(x_\beta) < v_\beta).$$

(ii) A family $\{x_\beta\}$ is said to be $\mathbf{v}-$*Cauchy* if

$$\forall\beta\forall\gamma(\beta < \gamma < \alpha \implies d(x_\beta, x_\gamma) < v_\beta).$$

(iii) A generalized metric space M is said to be *complete* if for every $\mathbf{v}$, every $\mathbf{v}-$Cauchy family $\mathbf{v}-$clusters to some element of M.
(iv) A set $A \subseteq M$ will be called *complete* if for every $\mathbf{v}$, whenever an $\mathbf{v}-$Cauchy family consists of elements of A, it $\mathbf{v}-$ cluster to some element of A.

Comments.
1. Clustering is similar to convergence, but not exactly. The main difference is that if a sequence converges to x, then this x is uniquely determined by this sequence, while a family can cluster to different points. E.g., if we take $v_n = 2^{-n}$, and $x_n = 2^{-n}$, then x_n clusters to both 0 and $2^{-\omega}$.

In the same sense, completeness of a subset A is similar to closeness.

2. For our set of values, only division by 2 makes sense (namely, $(1/2) \cdot 2^{-\alpha} = 2^{-(\alpha+1)}$), therefore, we can define contraction only for $k = 1/2$.

A2.4. Fixed point theorem for multi-valued mappings in a generalized metric space

Definition 8′. Let's say that a multi-values mapping $T : M \to 2^M$ is a $(1/2)-$*contraction* if for every $x \in M$, for every $y \in M$, and for every $a \in T(x)$, there exists a $b \in T(y)$ such that $d(a, b) \le (1/2)d(x, y)$.

FIXED POINT THEOREM FOR MULTI-VALUED CONTRACTIONS OF GENERALIZED METRIC SPACES. *Assume that M is a complete generalized metric space, T is a multi-valued $(1/2)-$contraction on M such that $T(x)$ is not empty for some $x \in M$ (i.e., T is not identically empty), and for every $x \in M$, the set $T(x)$ is complete. Then, T has a fixed point.*

Proof. Since T is not identically empty, there exists an element $x_0 \in M$ such that $T(x_0) \ne \phi$. Let's choose an arbitrary element $x_1 \in T(x_0)$. Since T is a $(1/2)-$contraction, there exists an $x_2 \in T(x_1)$ such that $d(x_1, x_2) \le (1/2)d(x_0, x_1)$. We can apply the same argument again, and thus step-by-step, we will contruct a sequence x_n such that for every n, we have $x_{n+1} \in T(x_n)$ and $d(x_{n+1}, x_{n+2}) \le (1/2)d(x_n, x_{n+1})$. Therefore, $d(x_n, x_{n+1}) \le 2^{-n}d(x_0, x_1)$. Since d is an ultra-metric (i.e., triangle inequality is true for max instead of $+$), we can conclude that for integers m and n, if $m < n$, then $d(x_m, x_n) \le 2^{-m}d(x_0, x_1)$.

Let us now extend (using transfinite recursion) this sequence into a family $\{x_\beta\}_{\beta<\omega_1}$ with the property that $x_{\beta+1} \in T(x_\beta)$, and if $\beta < \gamma$, then $d(x_\beta, x_\gamma) \le 2^{-\beta}d(x_0, x_1)$ (here, we define $2^{-\beta}2^{-\gamma}$ as $2^{-(\gamma+\beta)}$).

This transfinite recursion will run as follows. Assume that we have already described x_β for all $\beta < \alpha$. How to choose x_α?

This construction will be different in three different cases:

1) when an ordinal α has at least two preceding ones (i.e., when $\alpha = \gamma + 1$ for some γ);
2) when an ordinal α has only on preceding ordinal, i.e., when $\alpha = \gamma + 1$ for a limit ordinal γ;
3) when α is a limit ordinal itself.

1. Assume that $\alpha = \gamma + 2$. Then, we have x_γ and $x_{\gamma+1} \in T(x_\gamma)$ such that $d(x_\gamma, x_{\gamma+1}) \le 2^{-\gamma} d(x_0, x_1)$. Since T is a $(1/2)-$contraction, there exists an element $x_\alpha \in T(x_{\gamma+1})$ for which $d(x_\alpha, x_{\gamma+1}) \le (1/2)2^{-\gamma} d(x_0, x_1) = 2^{-(\gamma+1)} d(x_0, x_1)$. Using triangle inequality, one can easily check that when $\beta < \gamma + 1$, we have

$$d(x_\alpha, x_\beta) \le \max(d(x_\alpha, x_{\gamma+1}), d(x_{\gamma+1}, x_\beta)) = \max(2^{-(\gamma+1)}, 2^{-\beta}) = 2^\beta.$$

2. Suppose that x_β is constructed for all $\beta < \alpha$, and $\alpha = \gamma + 1$, where γ is a limit ordinal. Then, for every $\beta < \gamma$, from $d(x_\beta, x_\gamma) \le 2^{-\beta} d(x_0, x_1)$ and from the fact that T is a $(1/2)-$contraction, we conclude that there exists a $y_\beta \in T(x_\gamma)$ such that $d(y_\beta, x_{\beta+1}) \le 2^{-(\beta+1)} d(x_0, x_1)$.

From the ultrametric triangle inequality, we can now conclude that if $\beta < \delta$, then $d(y_\beta, y_\delta) \le 2^{-(\beta+1)} d(x_0, x_1)$. Therefore, for an appropriate $\mathbf{v}$ (namely, for $v_\beta = 2^{-(\beta+1)} d(x_0, x_1)$), the family y_β is a $\mathbf{v}-$Cauchy family. All values y_β belong to $T(x_\gamma)$. We assumed that $T(x)$ is complete for each x. Therefore, in $T(x_\gamma)$, there exists a value x_α to which the sequence y_β $\mathbf{v}-$clusters, i.e., for which $d(x_\alpha, y_\beta) \le 2^{-(\beta+1)} d(x_0, x_1)$.

From the ultrametric triangle inequality, we can now conclude that $d(x_\alpha, x_{\beta+1}) \le 2^{-(\beta+1)} d(x_0, x_1)$ for all $\beta < \gamma$.

From this inequality, we can conclude that $d(x_\gamma, x_\alpha) \le \max(x_\alpha, x_{\beta+1}), d(x_{\beta+1}, x_\gamma)) \le 2^{-(\beta+1)} d(x_0, x_1)$ for all $\beta < \gamma$. Since γ is a limit ordinal, we can thus conclude that $d(x_\gamma, x_\alpha) \le 2^{-\gamma} d(x_0, x_1)$.

3. Assume that α is a limit ordinal, and that x_β are already constructed for all $\beta < \alpha$. Then, the family $\{x_\beta\}_{\beta<\alpha}$ is $\mathbf{v}-$Cauchy for $v_\beta = 2^{-\beta} d(x_0, x_1)$. Since M is complete, there exists an $x \in M$ to which x_β $\mathbf{v}-$clusters, i.e., for which $d(x_\beta, x) \le 2^{-\beta} d(x_0, x_1)$ for all $\beta < \alpha$. So, we can take this x as the desired x_α.

In all three cases, we have constructed x_α with the desired property. Now, we have a sequence x_α for all $\alpha < \omega_1$ such that if $\alpha < \beta$, then $d(x_\alpha, x_\beta) \le 2^{-\alpha} d(x_0, x_1)$. So, we have a $\mathbf{v}-$Cauchy sequence, and since M is complete, it $\mathbf{v}-$clusters to some $x \in M$ for which $d(x_\alpha, x) \le 2^{-\alpha} d(x_0, x_1)$. Let us prove that this x is the desired fixed point.

Indeed, for every α, from the facts that $x_{\alpha+1} \in T(x_\alpha)$ and that T is a $(1/2)-$contraction, it follows that there exists a $y_\alpha \in T(x)$ such that $d(y_\alpha, x_{\alpha+1}) \le (1/2) d(x, y_\alpha) \le 2^{-(\alpha+1)} d(x_0, x_1)$. From triangle inequality, we can now conclude that if $\beta < \alpha$, then $d(y_\alpha, y_\beta) \le 2^{-\beta} d(x_0, x_1)$. So, y_β is a $\mathbf{v}-$Cauchy family of elements of $T(x)$. Since $T(x)$ is complete, there exists a $y \in T(x)$ to which y_α $\mathbf{v}-$clusters, i.e., for which $d(y, y_\alpha) \le 2^{-\alpha} d(x_0, x_1)$.

Therefore, for every countable ordinal β,

$$d(x,y) \le \max(d(x, x_{\beta+1}), d(x_{\beta+1}, y_\beta), d(y_\beta, y)) \le 2^{-\beta} d(x_0, x_1).$$

According to our definition of V, the only value from V that is smaller than all values $2^{-\beta}$ is 0. So, $d(x,y) = 0$, hence $x = y$. Since $y \in T(x)$, we thus conclude that $x = y \in T(x)$, i.e., that x is a fixed point of T. Q.E.D.

A3. THIS FIXED POINT THEOREM CAN BE APPLIED TO LOCALLY STRATIFIED DISJUNCTIVE LOGIC PROGRAMS

THEOREM. *Let* Π *be a locally stratified extended disjunctive logic program. Then* Π *has an answer set.*

Idea of the proof. Let Lit_α denote strata of Π. Define the generalized distance d on the set $M = 2^{Lit}$ as follows:

- if $A = B$, then $d(A, B) = 0$;
- if $A \ne B$, then $d(A, B) = 2^{-\alpha}$, where α is the smallest ordinal for which $A \cup Lit_\alpha \ne B \cup Lit_\alpha$.

Similarly to the proof of our Main Theorem, one can prove that thus defined generalized metric space is complete, and that the multi-valued mapping $T(S) = \alpha(\Pi^S)$ satisfies the assumptions of the fixed point theorem. Therefore, T has a fixed point which is an answer set to the program Π. Q.E.D.