Fuzzy Logic in Natural Language Processing

...wild speculation about the nature of truth, and other equally unscientific endeavors.

Richard Bergmair

Acknowledgments

thanks for supervising the project!

Ann Copestake

thanks for helping with the fuzzy logic!

Ulrich Bodenhofer

thanks for reading related manuscripts!

Ted Briscoe
Daniel Osherson

thanks for participating in the experiment!

MPhil students 05/06, NLIP Group,
RMRS-list, personal friends

Motivation

a small city near San Francisco

(Zadeh)

What does small'(x) mean in terms of population? What does near'(x,y) mean in terms of distance?

How do we deal with the vagueness involved in *small* and *near* ?

Natural Language Database Demo Interface

Query: hot dry city Submit

dof	mainid	x4.placeid	x4.placename	x4.type	x4.lat	x4.long	x4.pop	x4.temp	x4.wet
1.000	76	76	Blythe	city	984201	-1999974	8428	21	26
1.000	90	90	Brawley	city	995189	-2016437	18923	21	26
1.000	103	103	Calexico	city	1000449	-2015868	18633	20	27
1.000	106	106	Calipatria	city	992616	-2016162	2690	21	26
1.000	218	218	East Blythe	CDP	984161	-1999751	1511	21	26
1.000	233	233	El Centro	city	998552	-2016891	31384	21	26
1.000	326	326	Heber	CDP	999477	-2016206	2566	20	27
1.000	340	340	Holtville	city	998089	-2013714	4820	21	26
1.000	351	351	Imperial	city	997621	-2017094	4113	21	26
1.000	535	535	Niland	CDP	990674	-2016084	1183	21	26
1.000	727	727	Seeley	CDP	998519	-2019000	1228	21	26
1.000	842	842	Westmorland	city	994195	-2017975	1380	21	26
0.950	70	70	Big River	CDP	974939	-1995968	705	20	30
0.950	75	75	Bluewater	CDP	974337	-1994399	261	20	30
0.688	152	152	Coachella	city	982953	-2027239	16896	18	32
0.688	354	354	Indio	city	982274	-2028572	36793	18	32
0.688	483	483	Mecca	CDP	984786	-2025833	1966	18	32
0.625	529	529	Needles	city	963175	-2000378	5191	18	37
0.562	81	81	Bonita	CDP	1000641	-2042556	12542	17	33
0.562	114	114	Camp Pendleton South	CDP	990765	-2048554	11299	17	33
0.562	117	117	Carlsbad	city	992676	-2047024	63126	17	33

Outline

fuzzy logic as a generalization of bivalent logic

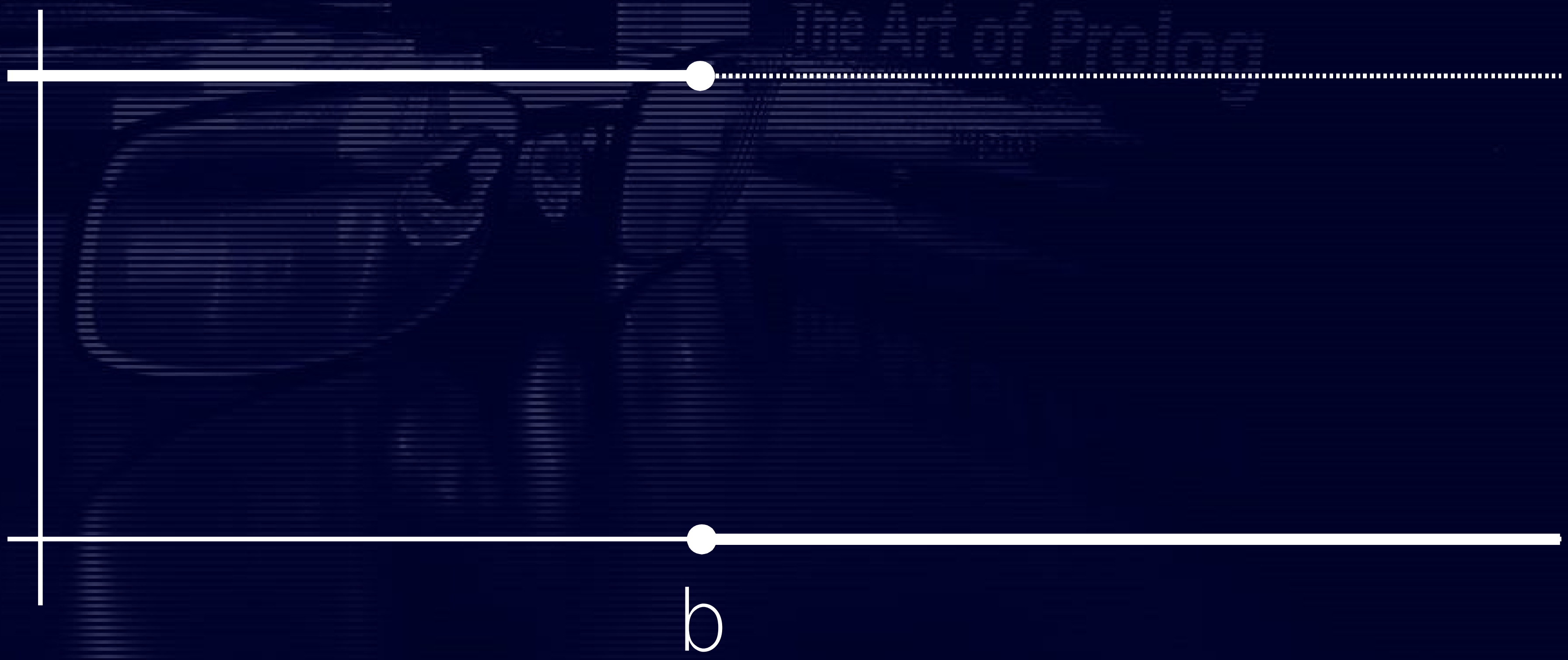
fuzzy logic in language modelling as a generalization of probabilistic models

vagueness and fuzzy semantics

putting fuzzy semantics to use in closed domain question answering

Bivalent Logic

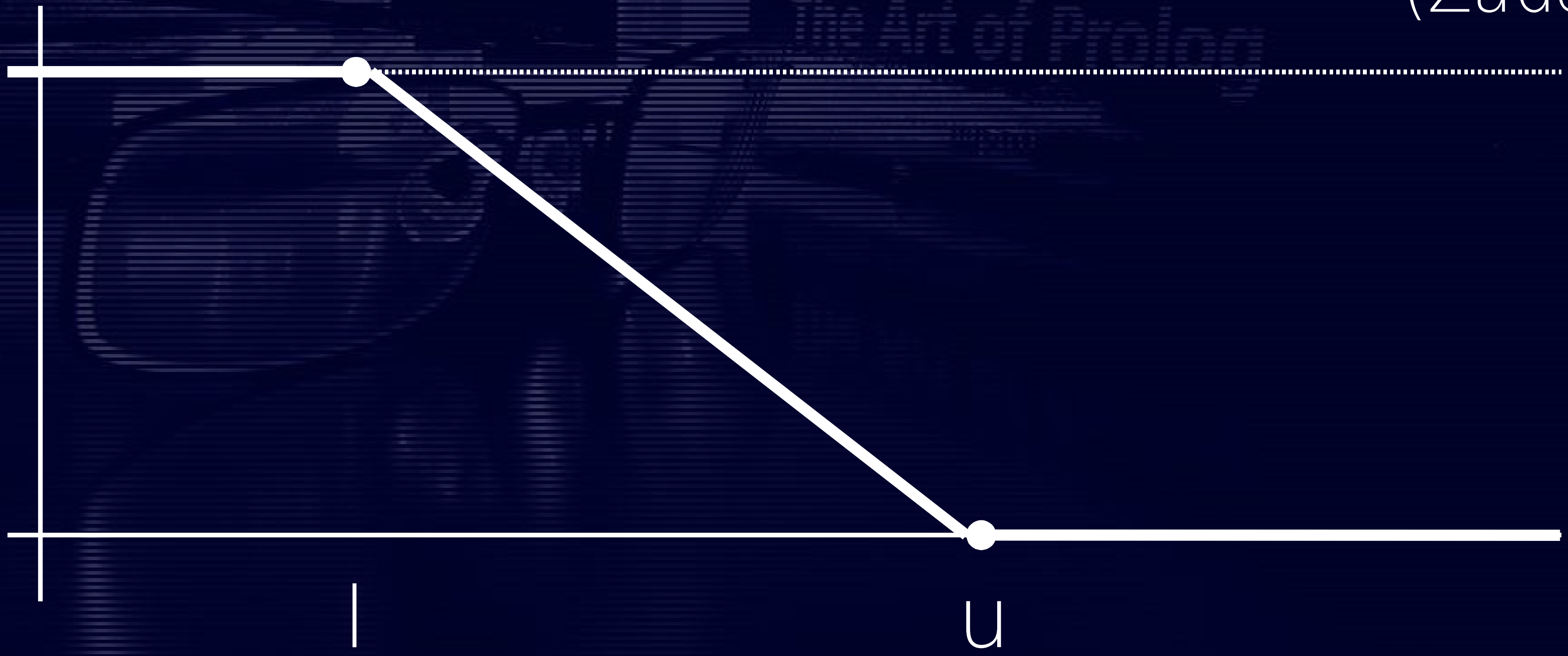
In classical logic: A is a set on domain X iff ∃ characteristic function $\chi_A : X \rightarrow \{0,1\}$ such that $\chi_A(x)=1$ iff $x \in A$



Fuzzy Logic

In fuzzy logic: A is a set on domain X iff ∃ characteristic function $\mu_A : X \rightarrow [0,1]$ such that $\mu_A(x)$ is a **degree** of membership.

(Zadeh)



Fuzzy Logic

Let A,B,C be fuzzy sets on X. Then $C = A \cap B$ with $\mu_C(x)=\mu_A(x)\wedge\mu_B(x)$ iff $\wedge:[0,1]\times[0,1]\rightarrow[0,1]$ with

(see Klement)

(1) $a\wedge b = b\wedge a$
(2) $a\wedge(b\wedge c) = (a\wedge b)\wedge c$
(3) $a\leq b \Rightarrow (a\wedge c)\leq(b\wedge c)$
(4) $a\wedge 1 = a$

These functions are known as *triangular norms.*

Fuzzy Logic

standard triangular norms:

$\wedge_M(x,y) = \min(x,y)$

$\wedge_P(x,y) = x*y$

$\wedge_L(x,y) = \max(x+y-1,0)$

$\wedge_D(x,y) = x$ if $y=1$, y if $x=1$, 0 othw.

Fuzzy Logic

Gödel logic is the logic induced by the minimum t-norm:

$$x \wedge y = \min(x, y)$$
$$x \vee y = \max(x, y)$$
$$\neg x = 1 - x$$

Fuzzy Logic

Product logic is the logic induced by the product t-norm:

$x \wedge y = x*y$

$x \vee y = x+y-x*y$

$\neg x = 1-x$

Fuzzy Logic

Łucasiewicz logic is the logic induced by the Łucasiewicz t-norm:

$x \wedge y = \max(x+y-1, 0)$

$x \vee y = \min(x+y, 1)$

$\neg x = 1-x$

Fuzzy Logic

More generally: Frank-family t-norms:

$$\wedge^{\lambda}_{F}(x,y) := \log_{\lambda}\left(1+ \frac{(\lambda^{x}-1)(\lambda^{y}-1)}{\lambda-1}\right)$$

$$\wedge^{0}_{F} := \wedge_{M},\ \wedge^{1}_{F} := \wedge_{P},\ \wedge^{\infty} := \wedge_{L}$$

Schweizer-Sklar-family t-norms:

$$\wedge^{\lambda}_{SS}(x,y) := (\max(x^{\lambda}+y^{\lambda}-1,0))^{1/\lambda}$$

$$\wedge^{-\infty}_{SS} := \wedge_{M},\ \wedge^{0}_{SS} := \wedge_{P},\ \wedge^{\infty} := \wedge_{D}$$

Outline

fuzzy logic as a generalization of bivalent logic

fuzzy logic in language modelling as a generalization of probabilistic models

vagueness and fuzzy semantics

putting fuzzy semantics to use in closed domain question answering

Fuzzy N-grams, regular lg.

fuzzy n-grams

$$\mu_L(\langle x_1,\ldots,x_K\rangle)=\bigwedge_{i=1}^{K-N} \mu(x_i,x_{i+1},x_{i+N})$$

fuzzy regular languages (Gaines & Kohout, Doostfatemeh et al, etc.)

$$\mu_L(\langle x_1,\ldots,x_K\rangle)=\bigvee_{S} \bigwedge_{i=1}^{K-1} \mu_\delta(s(i),s(i+1))\wedge\mu_{s(i+1)}(x_{i+1})$$

Fuzzy context-free lg.

fuzzy context-free languages

$$\mu_L(\langle x_1,\ldots,x_J\rangle) = \bigvee_{\langle d_1,\ldots,d_K\rangle} \bigwedge_{i=1}^{K} \mu(d_i, C(\langle d_1,\ldots,d_i\rangle))$$

(Lee & Zadeh,
Carter et al.)

...and so on, up the Chomsky hierarchy.

Fuzzy Language Models

Well this is a nice generalization…

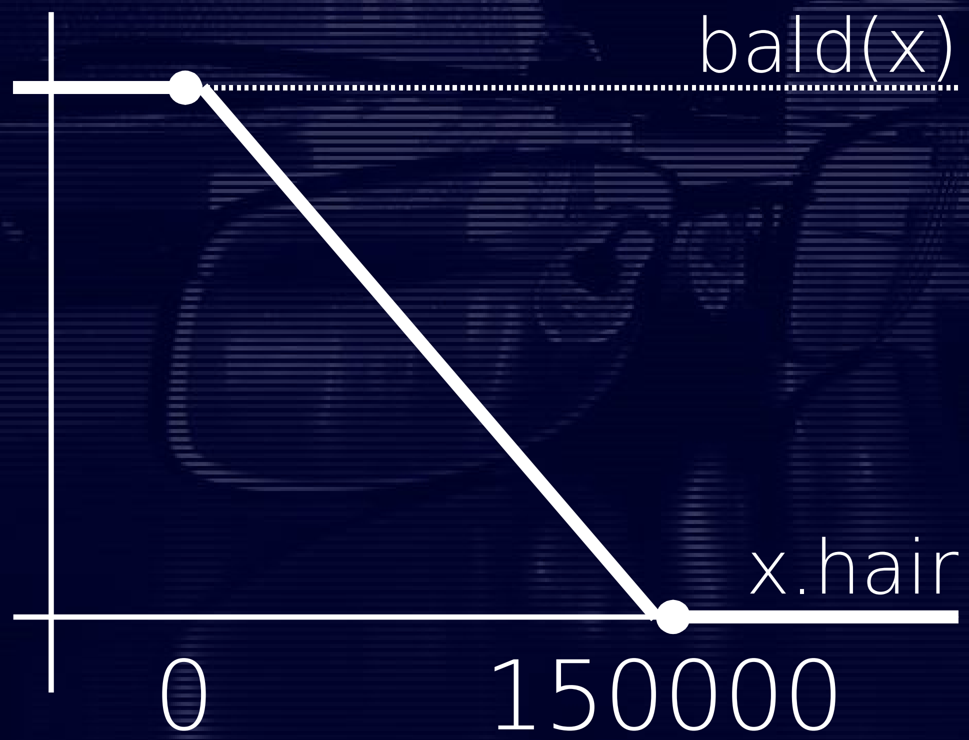
…but is there a linguistic reality to this? …

Work on inducing FCFGs from the SUSANNE corpus by Carter et. al (disappointing results)

…for syntax I don't see one.

Fuzzy Semantics

...for semantics, denotations are hard to define using probability densities.



x.hair = 76273
bald(x) = ?

Fuzzy Semantics

…and independence assumptions are difficult to justify.

Syntax: independence holds!

l_1:cold(x_1), l_2:rainy(x_2), l_3:town(x_3)

$l_1=l_2, l_2=l_3, x_1=x_2, x_2=x_3$

Semantics: independence does not hold!

l_1:cold(x_1), l_1:rainy(x_1), l_1:town(x_1)

Outline

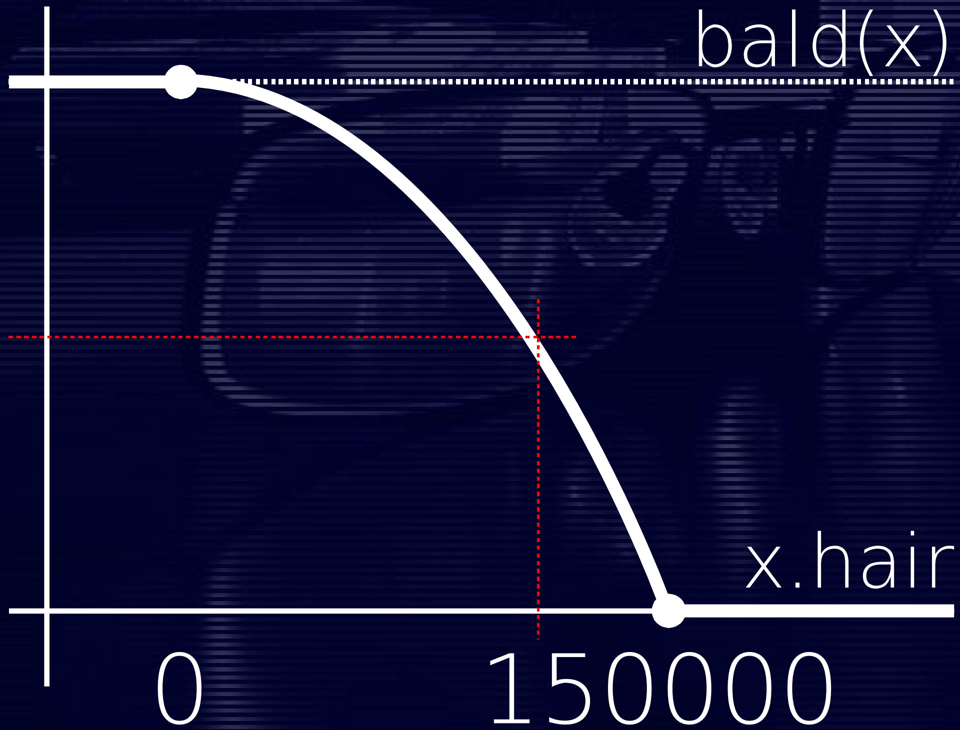
fuzzy logic as a generalization of bivalent logic

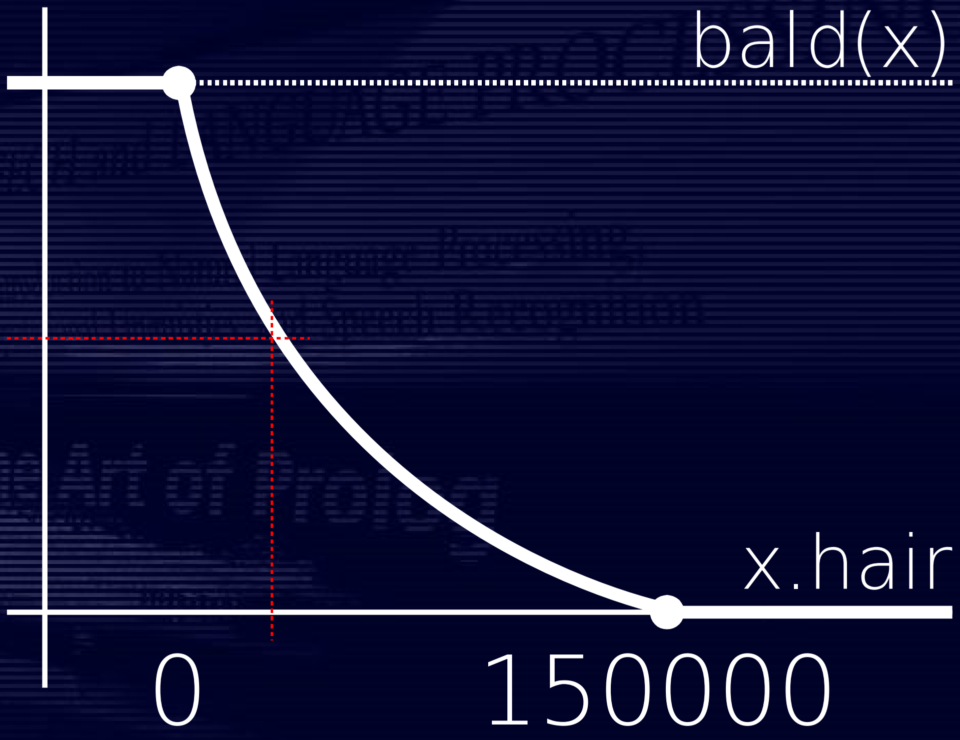
fuzzy logic in language modelling as a generalization of probabilistic models

vagueness and fuzzy semantics

putting fuzzy semantics to use in closed domain question answering

Fuzzy Semantics





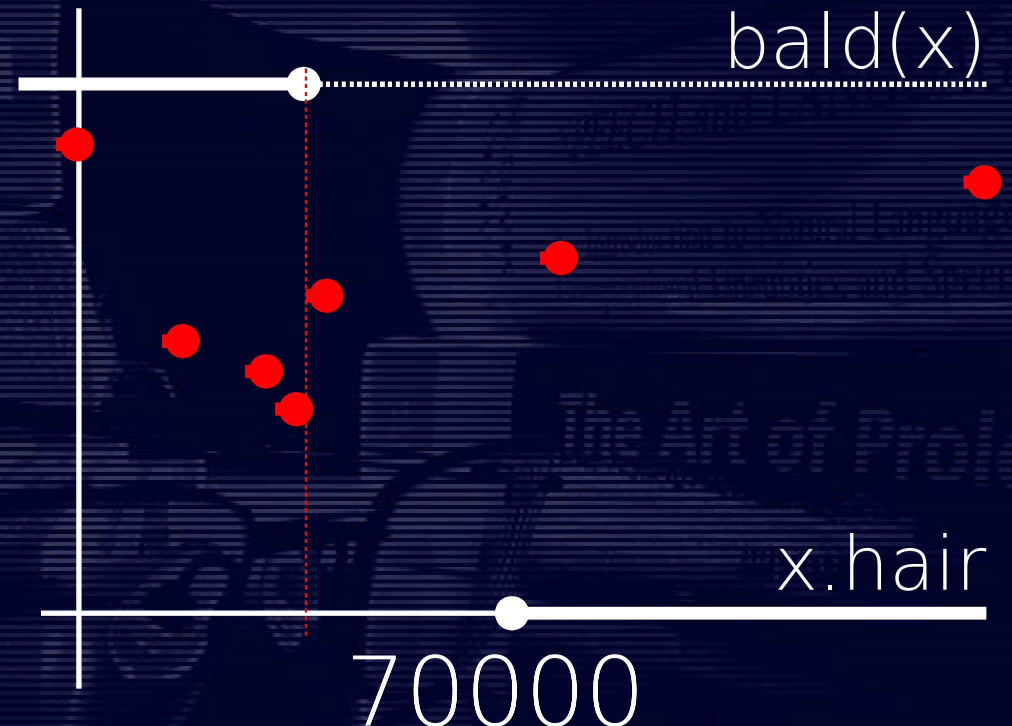
Fuzzy Semantics Experiment

*If a city had a year-round average temperature of **12 degrees celsius**, it would be natural to call it a cold city: (yes/no)*

If a skyscraper had 78 floors it would be natural to call it a rather tall skyscraper: (yes/no)

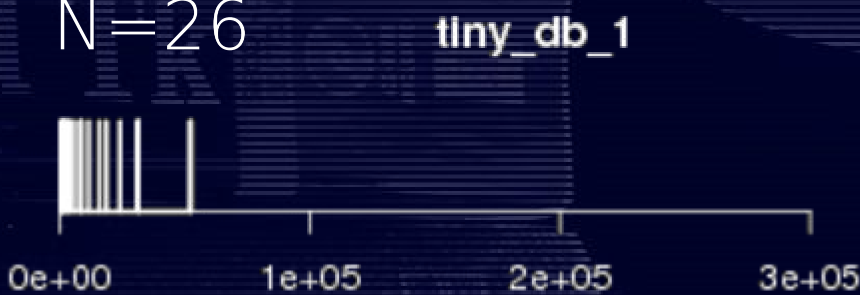
...

Fuzzy Semantics Experiment



cities domain

N=26

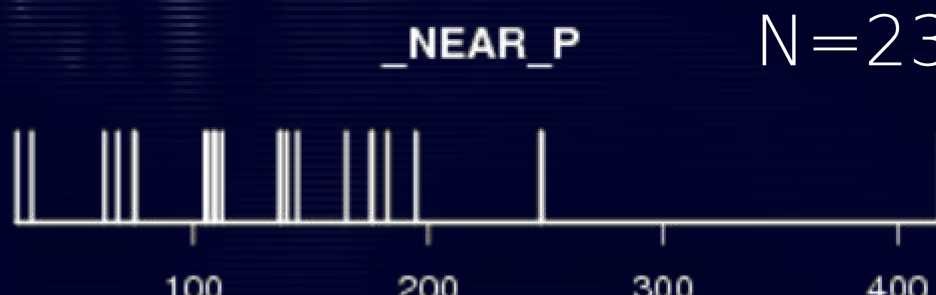


N=26

N=25

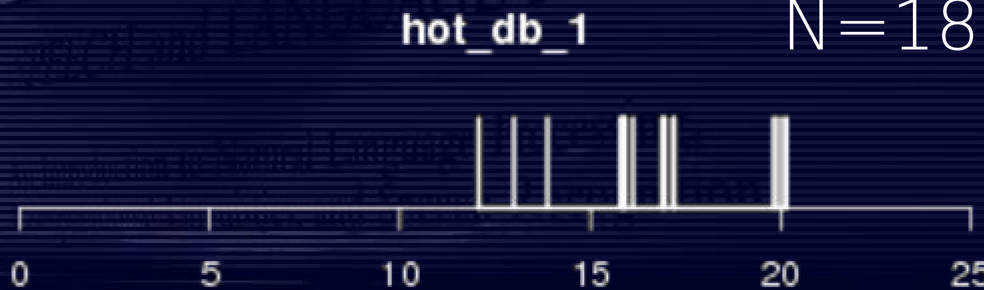
N=26

N=23



cities domain (cont'd)

N=18



N=18

N=13

N=13

skyscrapers domain

Fuzzy Semantics Experiment

What does this tell us about
Fuzzy Semantics?

1. Membership can clearly be judged as nonincreasing or nondecreasing.

...consistent with the observations about most predicates – but not all due to mistakes in the experimental setup.

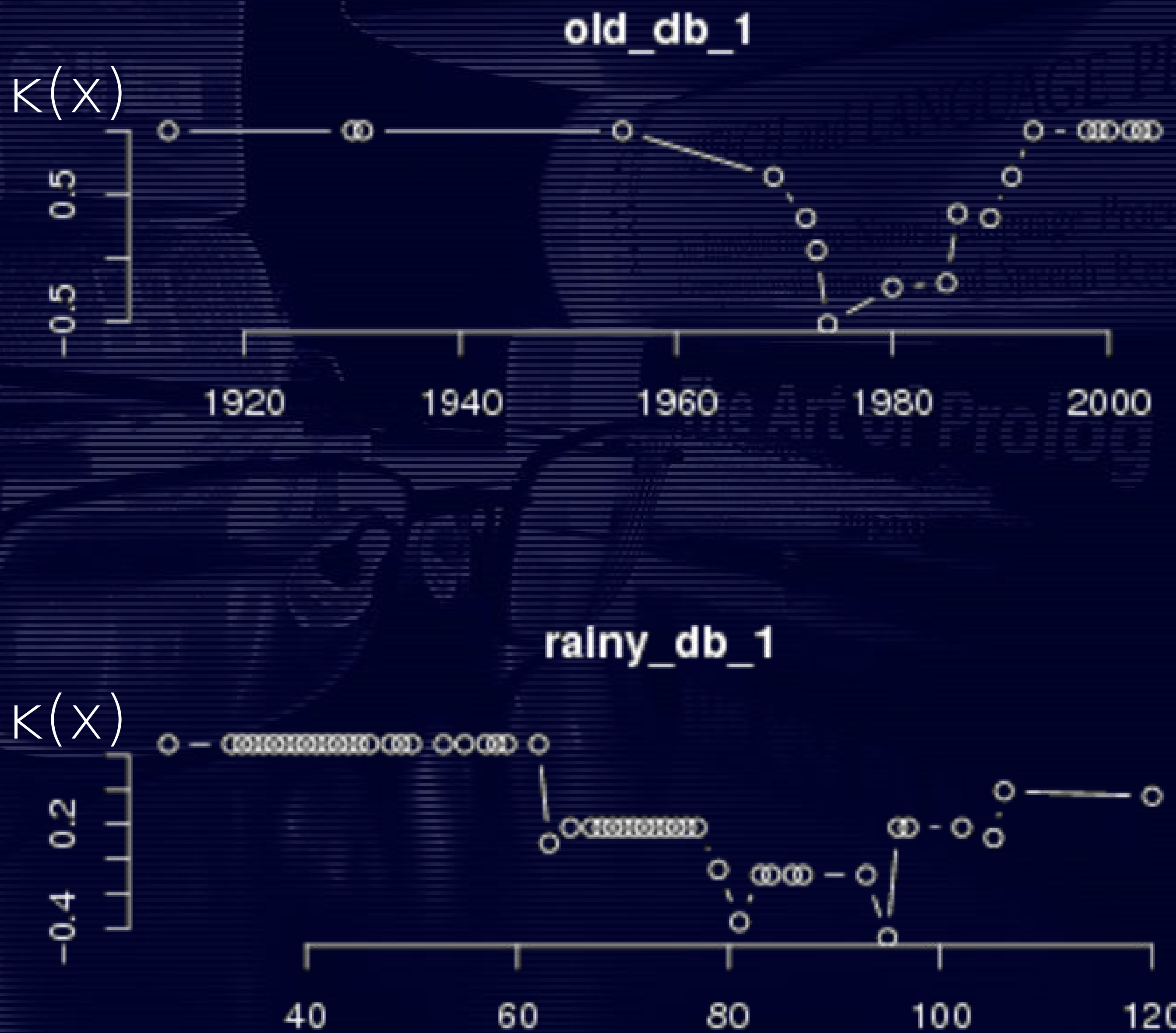
Fuzzy Semantics Experiment

What does this tell us about Fuzzy Semantics?

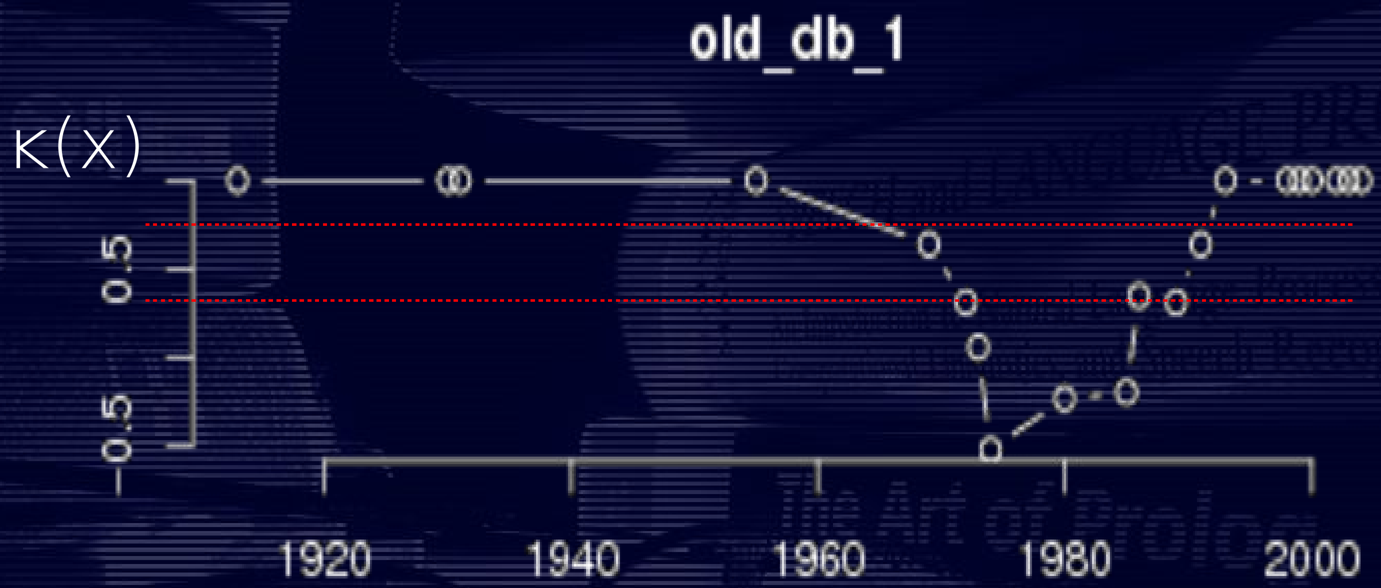
2. A “region of fuzzy membership” can always be clearly identified and distinguished from a region of crisp membership.

…turned out to be tricky to test.

Fuzzy Semantics Experiment



Fuzzy Semantics Experiment



Fuzzy Semantics Experiment

2. A “region of fuzzy membership” can always be clearly identified and distinguished from a region of crisp membership.

…consistent with the observations about most predicates – but not all due to mistakes in the experimental setup.

Fuzzy Semantics Experiment

What does this tell us about
Fuzzy Semantics?

3. Decision boundaries as well as fuzzy sets may be contradictory across speakers, but are always consistent for each speaker in isolation.

Clearly consistent with observations!

Ordering-based Semantics

Outline

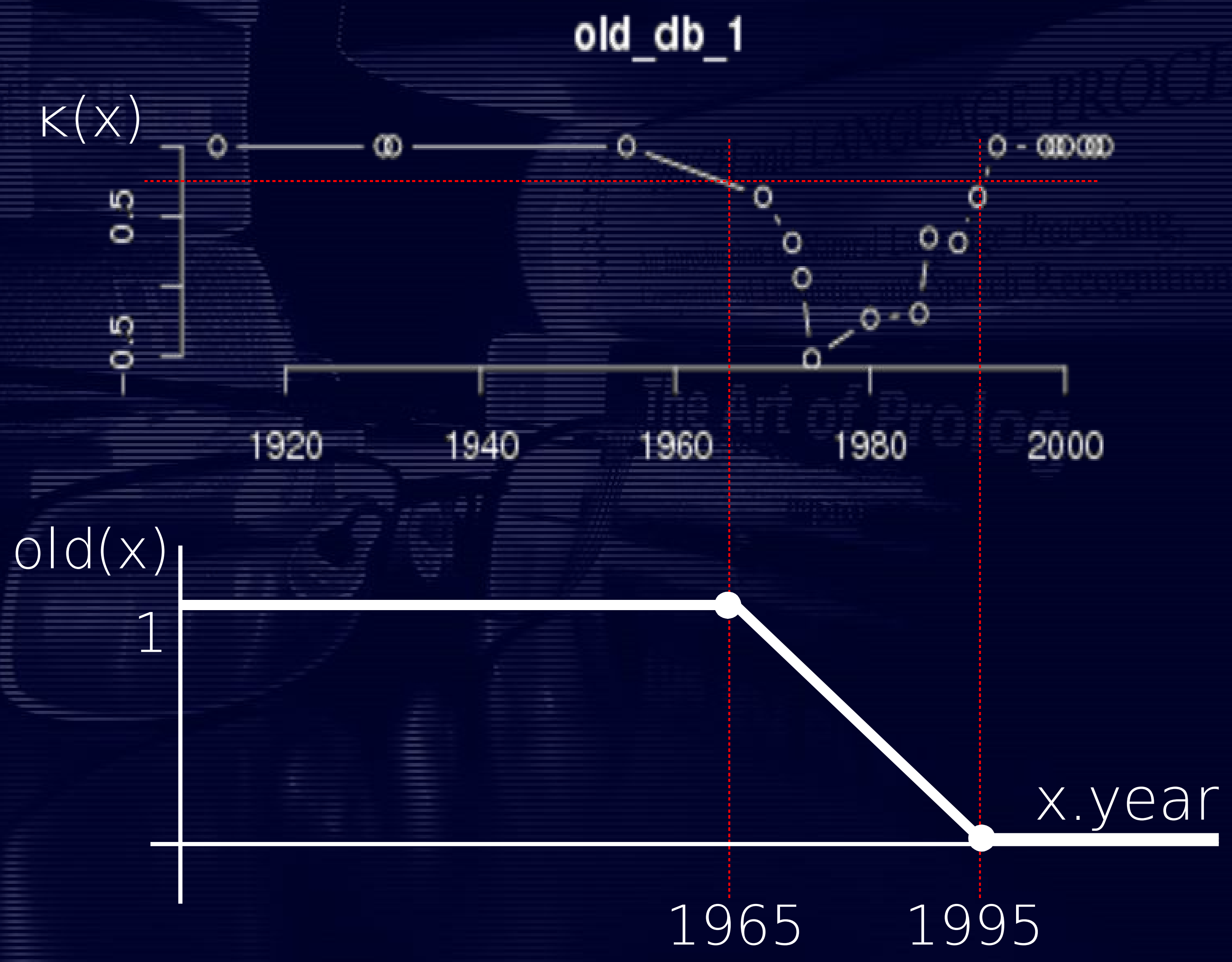
fuzzy logic as a generalization of bivalent logic

fuzzy logic in language modelling as a generalization of probabilistic models

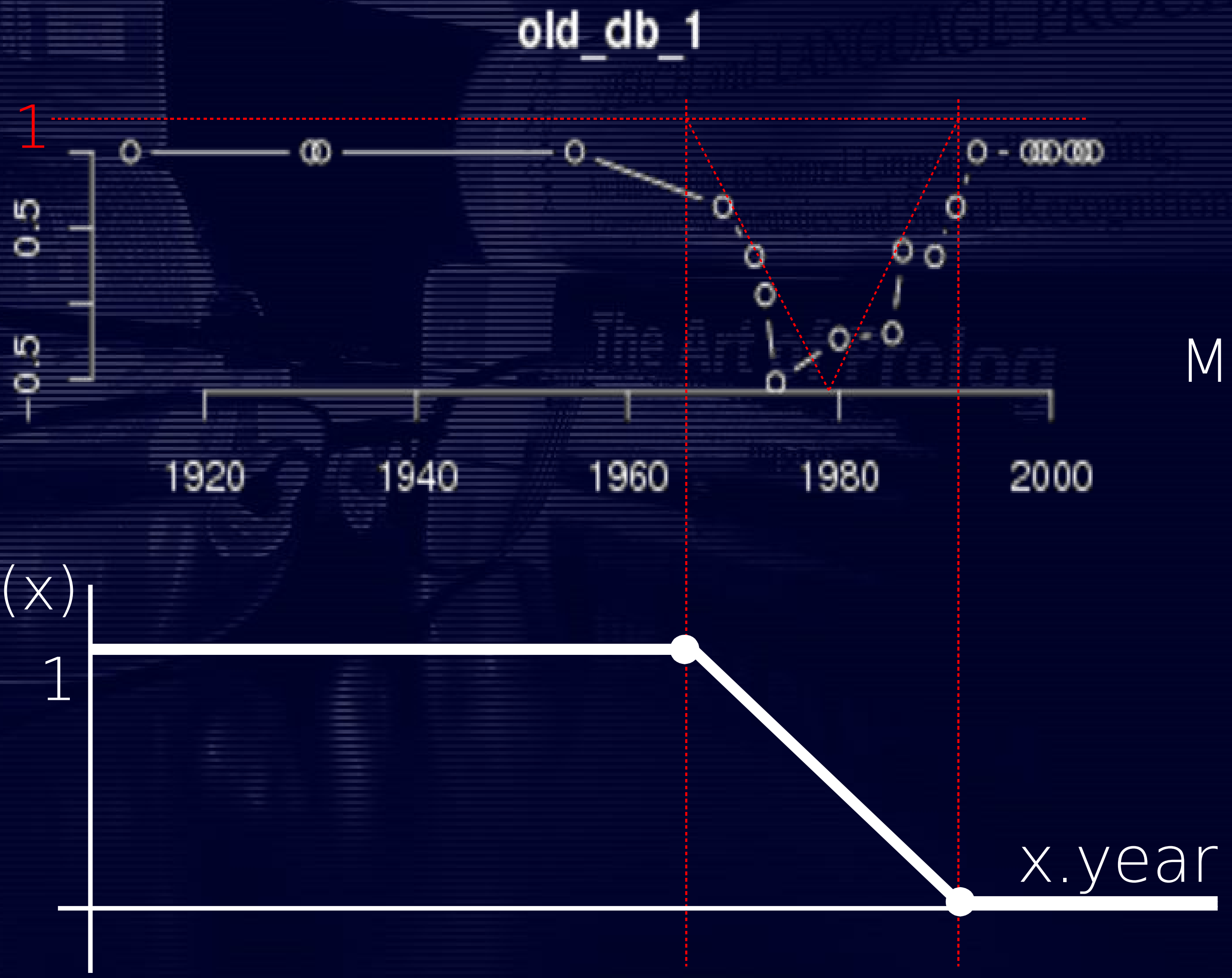
vagueness and fuzzy semantics

putting fuzzy semantics to use in closed domain question answering

Characteristic Functions



Characteristic Functions



MSE=.16

Database Interface

hot dry city

```
SELECT
  x.*,
  hot(x.temp)∧dry(x.rainfall) AS mu
FROM
  place
WHERE
  mu > 0
ORDER BY mu DESC
```

Database Interface

small city near San Francisco

```
SELECT
    x.*, z.*, y.*,
    small(x)∧near(z) AS mu
FROM
  place x,
  refnear z,
  place y
WHERE
  x.placeid = z.placeid AND
  z.fkplaceid = y.placeid AND
  y.name = 'San Francisco' AND
  mu > 0
ORDER BY mu DESC
```

Database Interface

dry city near a rainy city

```
SELECT
    x.*, z.*, y.*,
    dry(x)∧near(z)∧rainy(y) AS mu
FROM
   place x,
   refnear z,
   place y
WHERE
   x.placeid = z.placeid AND
   z.fkplaceid = y.placeid AND
   mu > 0
ORDER BY mu DESC
```

Linguistic Data Modelling

```
LEXENT adv {
  STEM "rather";
  TYPE "adv_degree_spec_le";
};

ENTITY place {

  LEXENT noun {
    STEM "city";
    TYPE "n_intr_le";
    ONSET "con";
  };

  PK placeid;

  GEN nb "#noun";

  INTAT lat;
  INTAT long;
```

```
INTAT temp {
    LEXENT adj {
      STEM "hot";
      TYPE "adj_intrans_le";
      ONSET "con";
    };
   LEXENT adj {
      STEM "cold";
      TYPE "adj_intrans_le";
      ONSET "con";
    };
    GEN ap "#adv #adj";
    GEN nb "#ap #noun";
    DSCR "If a city had ayear-round
      average <B>temperature of
      #temp</B> degrees celsius, it
      would be natural to call it a
      <B>#ap</B> city.";
  };
};
```

Linguistic Data Modelling

```
ENTITY place {
  ...
  INTAT temp { ... };
  STRAT(10) type;
  ID(100) placename {
    TYPE "n_proper_city_le";
    ONSET "con";
  };
  REFERENCE refnear TO MANY place {
    INTAT distance {
      LEXENT near {
        STEM "near";
          TYPE "p_reg_le";
          ONSET "con";
          REL "_NEAR_P_REL";
        };
      DSCR "If a city was a distance ..."
    };
  };
}
```

Outline

fuzzy logic as a generalization of bivalent logic

fuzzy logic in language modelling as a generalization of probabilistic models

vagueness and fuzzy semantics

putting fuzzy semantics to use in closed domain question answering