## Prefix-vocabulary classes

$$
\left[\Pi,\left(p_{1}, p_{2}, \ldots\right),\left(f_{1}, f_{2}, \ldots\right)\right]_{(=)}
$$

- $\quad \Pi$ is a word over $\left\{\exists, \forall, \exists^{*}, \forall^{*}\right\}$, describing set of quantifier prefixes
- $\quad p_{m}, f_{m} \leq \omega$ indicate how many relation and function symbols of arity $m$ may occur
- presence or absence of $=$ indicates whether the formulae may contain equality

Example: $\quad\left[\exists^{*} \forall \exists^{*},(\omega, 1) \text {, all }\right]_{=}$ sentences $\exists x_{1} \ldots \exists x_{m} \forall y \exists z_{1} \ldots \exists z_{n} \varphi$ where $\varphi$ is quantifier-free and

- contains at most one binary predicate, and no predicates of arity $\geq 3$,
- may contain any number of monadic predicates,
- may contain any number of function symbols of any arity,
- may contain equality.


## The complete classification: undecidable cases

A: Pure predicate logic (without functions, without =)

| (1) | $[\forall \exists \forall,(\omega, 1),(0)]$ | (Kahr 1962) |
| :--- | :--- | :--- |
| (2) | $\left[\forall^{3} \exists,(\omega, 1),(0)\right]$ | (Surányi 1959) |
| (3) | $\left[\forall^{*} \exists,(0,1),(0)\right]$ | (Kalmár-Surányi 1950) |
| $(4)$ | $\left[\forall \exists \forall^{*},(0,1),(0)\right]$ | (Denton 1963) |
| $(5)$ | $\left[\forall \exists \forall \exists^{*},(0,1),(0)\right]$ | (Gurevich 1966) |
| $(6)$ | $\left[\forall^{3} \exists^{*},(0,1),(0)\right]$ | (Kalmár-Surányi 1947) |
| $(7)$ | $\left[\forall \exists^{*} \forall,(0,1),(0)\right]$ | (Kostyrko-Genenz 1964) |
| $(8)$ | $\left[\exists^{*} \forall \exists \forall,(0,1),(0)\right]$ | (Surányi 1959) |
| $(9)$ | $\left[\exists^{*} \forall^{3} \exists,(0,1),(0)\right]$ | (Surányi 1959) |

## The complete classification: undecidable cases

B: Classes with functions or equality

| $(10)$ | $[\forall,(0),(2)]=$ | (Gurevich 1976) |
| :--- | :--- | :--- |
| $(11)$ | $[\forall,(0),(0,1)]_{=}$ | (Gurevich 1976) |
| $(12)$ | $\left[\forall^{2},(0,1),(1)\right]$ | (Gurevich 1969) |
| $(13)$ | $\left[\forall^{2},(1),(0,1)\right]$ | (Gurevich 1969) |
| $(14)$ | $\left[\forall^{2} \exists,(\omega, 1),(0)\right]_{=}$ | (Goldfarb 1984) |
| $(15)$ | $\left[\exists^{*} \forall^{2} \exists,(0,1),(0)\right]_{=}$ | (Goldfarb 1984) |
| $(16)$ | $\left[\forall^{2} \exists^{*},(0,1),(0)\right]_{=}$ | (Goldfarb 1984) |

## The complete classification: decidable cases

(Exclude the trivial classes: finite prefix and finite relational vocabulary)
A: Classes with the finite model property
(1) $\left[\exists^{*} \forall^{*}, \text { all, (0) }\right]_{=} \quad$ (Bernays, Schönfinkel 1928)
(2) $\left[\exists^{*} \forall^{2} \exists^{*}\right.$, all, (0)] (Gödel 1932, Kalmár 1933, Schütte 1934)
$\begin{array}{lll}\text { (3) } & {[\text { all, }(\omega),(\omega)]} & \text { (Löb 1967, Gure } \\ \text { (4) } & {\left[\exists^{*} \forall \exists^{*} \text {, all, all }\right]} & \text { (Gurevich 1973) }\end{array}$
(Gurevich 1976)

B: Classes with infinity axioms
(6) $[\text { all, }(\omega),(1)]_{=}$
(Rabin 1969)
(7)
$\left[\exists^{*} \forall \exists^{*}\right.$, all, (1) $\rfloor \leftrightarrows$ (Shelah 1977)

## $\mathcal{F} \mathcal{O}^{1}$ WITH COUNTING, $\mathcal{C}^{1}$

$\mathcal{C}^{1}$ : extension of $\mathcal{F} \mathcal{O}^{1}$ with counting quantifiers: $\exists \leq m, \exists \geq m, \exists=m$ meaning that there exists at most, at least, exactly $m$ elements satisfying some property.
$\exists \geq{ }^{125} x \top \wedge \exists^{=50} x$ French $(x) \wedge \exists \exists^{=36} x \operatorname{German}(x) \wedge \exists^{=36} x$ Spanish $(x)$
SATISFIABLE


## $\mathcal{F} \mathcal{O}^{1}$ WITH COUNTING, $\mathcal{C}^{1}$

$\mathcal{C}^{1}$ : extension of $\mathcal{F} \mathcal{O}^{1}$ with counting quantifiers: $\exists \leq m, \exists \geq m, \exists=m$ meaning that there exists at most, at least, exactly $m$ elements satisfying some property.

# $\exists \geq{ }^{125} x \top \wedge \exists^{=50} x$ French $(x) \wedge \exists^{=36} x \operatorname{German}(x) \wedge \exists^{=36} x \operatorname{Spanish}(x) \wedge$ $\forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x))$ <br> NOT SATISFIABLE 

$\exists=122 x \top \wedge \exists=50 x$ French $(x) \wedge \exists=36 x \operatorname{German}(x) \wedge \exists=36 x$ Spanish $(x)$


## $\mathcal{F} \mathcal{O}^{1}$ WITH COUNTING, $\mathcal{C}^{1}$

$\mathcal{C}^{1}$ : extension of $\mathcal{F} \mathcal{O}^{1}$ with counting quantifiers: $\exists \leq m, \exists \geq m, \exists=m$ meaning that there exists at most, at least, exactly $m$ elements satisfying some property.

$$
\begin{gathered}
\exists \geq 125 x \top \wedge \exists=50 x \text { French }(x) \wedge \exists^{=36} x \operatorname{German}(x) \wedge \exists=36 x \operatorname{Spanish}(x) \wedge \\
\forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x)) \\
\text { NOT SATISFIABLE }
\end{gathered}
$$

$$
\begin{gathered}
\exists=122 x \top \wedge \exists^{=50} x \operatorname{French}(x) \wedge \exists^{=36} x \operatorname{German}(x) \wedge \exists^{=36} x \operatorname{Spanish}(x) \wedge \\
\\
\forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x)) \wedge
\end{gathered}
$$

SATISFIABLE!

## $\mathcal{F} \mathcal{O}^{1}$ WITH COUNTING, $\mathcal{C}^{1}$

$\mathcal{C}^{1}$ : extension of $\mathcal{F} \mathcal{O}^{1}$ with counting quantifiers: $\exists \leq m, \exists \geq m, \exists=m$ meaning that there exists at most, at least, exactly $m$ elements satisfying some property.

$$
\begin{gathered}
\exists \geq 125 x \top \wedge \exists=50 x \text { French }(x) \wedge \exists^{=36} x \operatorname{German}(x) \wedge \exists=36 x \operatorname{Spanish}(x) \wedge \\
\forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x))
\end{gathered}
$$

## NOT SATISFIABLE

$$
\begin{array}{rl}
\exists=122 & x \top \wedge \\
& \exists=50 x \operatorname{French}(x) \wedge \exists=36 x \operatorname{German}(x) \wedge \exists=36 x \operatorname{Spanish}(x) \wedge \\
& \forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x)) \wedge \\
& \exists={ }^{=38} x(\operatorname{French}(x) \wedge \neg \operatorname{German}(x)) \wedge \\
& \exists=18 x(\operatorname{French}(x) \wedge \operatorname{Spanish}(x)) \wedge \\
& \exists=21 x(\operatorname{German}(x) \wedge \operatorname{Spanish}(x)) \wedge \\
& \exists={ }^{10} x(\operatorname{French}(x) \wedge \operatorname{German}(x) \wedge \operatorname{Spanish}(x))
\end{array}
$$

STILL SATISFIABLE ???

## Lemma (Normal form for $\mathcal{C}^{1}$ )

For every $\mathcal{C}^{1}$ formula $\varphi$ we can compute in polynomial time a formula
$\varphi^{\prime}$ of the form

$$
\varphi^{\prime}:=\bigwedge_{i=1}^{m} \exists \bowtie_{i} C_{i} x \varphi_{i}(x)
$$

satisfiable over the same domains as $\varphi$, where:

- $1 \leq m \leq|\varphi|$,
- each $\varphi_{i}$ is quantifier free,
- each $\bowtie_{i}$ is any of the symbols $\leq, \geq$ or $=$, and
- the $C_{i}$ are either one or occur as a quantifier subscript in $\varphi$.

Proof: similarly to $\mathcal{F} \mathcal{O}^{1}$ we replace subformulas of the form $\exists \bowtie \subset x \chi(x)$ with $\chi(x)$-quantifier-free, by new predicate symbols and add appropriate definitions.

## Theorem (FMP for $\mathcal{C}^{1}$ )

Let $\varphi$ be a formula in $\mathcal{C}^{1}$. If $\varphi$ is satisfiable, then it is satisfiable over a domain of size at most $2^{|\varphi|}$.

## Proof. WARNING! $\varphi=\exists_{x}^{22} T$

By the normal form Lemma we may assume that $\varphi$ has the form

$$
\varphi:=\bigwedge_{i=1}^{m} \exists \geq C_{i} x \theta_{i} \wedge \bigwedge_{j=1}^{m^{\prime}} \exists \leq D_{i} x \chi_{j} .
$$

Let $\mathfrak{A} \models \varphi$. For all $i(1 \leq i \leq m)$ select distinct elements $a_{i, 1}, \ldots, a_{i, c_{i}} \in A$ satisfying $\theta_{i}$ in $\mathfrak{A}$.
Let $B=\left\{a_{i, k} \mid 1 \leq i \leq m, 1 \leq k \leq C_{i}\right\}$, and let $\mathfrak{B}$ be the restriction of $\mathfrak{A}$ to $B$. Then $\mathfrak{B} \models \varphi$.

Corollary
SAT( $\left.\mathcal{C}^{1}\right)$ is in NExpTime.
Goal: NP

## COMPLEXITY OF $\mathcal{C}^{1}$

Our aim is to prove
Theorem
SAT( $\mathcal{C}^{1}$ ) is NP-complete.

$$
t_{p_{1}}^{4}(a)=\{B, R\}
$$

We cannot improve the bound on the size of minimal models: the formula $\exists^{\geq n} x P x$ has only models of exponential size with respect to $|\varphi|$.

Definition

$$
\left.t_{P_{1}}^{A /}(a)=B(x) \wedge R(x) \wedge\right\urcorner G(x)
$$

A 1-type of an element $a$ in a model $\mathfrak{A}$ is the conjunction of all literals satisfied by $a . \quad \Sigma=\{B, R, G\} \quad ๑^{B, R, \neg G}$
Idea: with each normal form $\varphi$ we associate a system of linear inequalities $\mathcal{E}_{\varphi}$ describing constraints on the number of distinct 1 -types realized in some model of $\varphi$.

## Systems of inequalities - Example

$\varphi:=\exists \exists^{=122} x \uparrow \wedge \exists^{=50} x \operatorname{French}(x) \wedge \exists^{=36} x \operatorname{German}(x) \wedge \exists^{=36} x \operatorname{Spanish}(x) \wedge$ $\forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x)) \wedge$
$\exists^{=38} x(\operatorname{French}(x) \wedge \neg \operatorname{German}(x)) \wedge$
$\exists^{=18} x(\operatorname{French}(x) \wedge \operatorname{Spanish}(x)) \wedge$
$\exists^{=21} x(\operatorname{German}(x) \wedge \operatorname{Spanish}(x)) \wedge$
$\exists=10 x(\operatorname{French}(x) \wedge \operatorname{German}(x) \wedge \operatorname{Spanish}(x))$
Denote the 1-types over the signature French, German, Spanish by $t_{\emptyset}, t_{F}, t_{G}, t_{S}, t_{F G}, t_{F S}, t_{G S}, t_{F G S}$ (the letters in the subscript indicate the positive subformulas of the type). $\mathcal{E}_{\varphi}$ contains:

## Systems of inequalities - EXAMPLE

$\varphi:=\exists \exists^{=122} x \upharpoonleft \wedge \exists^{=50} x \operatorname{French}(x) \wedge \exists^{=36} x \operatorname{German}(x) \wedge \exists \exists^{36} x \operatorname{Spanish}(x) \wedge$ $\forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x)) \wedge$ $\exists^{=38} x(\operatorname{French}(x) \wedge \neg \operatorname{German}(x)) \wedge$ $\exists^{=18} x(\operatorname{French}(x) \wedge \operatorname{Spanish}(x)) \wedge$ $\exists=21 x(\operatorname{German}(x) \wedge \operatorname{Spanish}(x)) \wedge$ $\exists=10 x(\operatorname{French}(x) \wedge \operatorname{German}(x) \wedge \operatorname{Spanish}(x))$
Denote the 1-types over the signature French, German, Spanish by $t_{\emptyset}, t_{F}, t_{G}, t_{S}, t_{F G}, t_{F S}, t_{G S}, t_{F G S}$ (the letters in the subscript indicate the positive subformulas of the type). $\mathcal{E}_{\varphi}$ contains:

$$
x_{\emptyset}+x_{F}+x_{G}+x_{S}+x_{F G}+x_{F S}+x_{G S}+x_{F G S}=122
$$

## Systems of inequalities - Example

$\varphi:=\exists \exists^{=122} x \uparrow \wedge \exists^{=50} x \operatorname{French}(x) \wedge \exists^{=36} x \operatorname{German}(x) \wedge \exists^{=36} x \operatorname{Spanish}(x) \wedge$ $\forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x)) \wedge$ $\exists=38 x(\operatorname{French}(x) \wedge \neg \operatorname{German}(x)) \wedge$ $\exists^{=18} x(\operatorname{French}(x) \wedge \operatorname{Spanish}(x)) \wedge$ $\exists^{=21} x(\operatorname{German}(x) \wedge \operatorname{Spanish}(x)) \wedge$ $\exists={ }^{10} x(\operatorname{French}(x) \wedge \operatorname{German}(x) \wedge \operatorname{Spanish}(x))$
Denote the 1-types over the signature French, German, Spanish by $t_{\emptyset}, t_{F}, t_{G}, t_{S}, t_{F G}, t_{F S}, t_{G S}, t_{F G S}$ (the letters in the subscript indicate the positive subformulas of the type). $\mathcal{E}_{\varphi}$ contains:

$$
\begin{aligned}
x_{\emptyset}+x_{F}+x_{G}+x_{S}+x_{F G}+x_{F S}+x_{G S}+x_{F G S} & =122 \\
x_{F}+x_{F G}+x_{F S}+x_{F G S} & =50
\end{aligned}
$$

## Systems of inequalities - Example

$$
\begin{aligned}
\varphi:=\exists=122 x \top \wedge & \exists=50 x \operatorname{French}(x) \wedge \exists=36 x \operatorname{German}(x) \wedge \exists=36 x \operatorname{Spanish}(x) \wedge \\
& \forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x)) \wedge \\
& \exists=38 x(\operatorname{French}(x) \wedge \neg \operatorname{German}(x)) \wedge \\
& \exists=18 x(\operatorname{French}(x) \wedge \operatorname{Spanish}(x)) \wedge \\
& \exists=21 x(\operatorname{German}(x) \wedge \operatorname{Spanish}(x)) \wedge \\
& \exists=10 x(\operatorname{French}(x) \wedge \operatorname{German}(x) \wedge \operatorname{Spanish}(x))
\end{aligned}
$$

Denote the 1-types over the signature French, German, Spanish by $t_{\emptyset}, t_{F}, t_{G}, t_{S}, t_{F G}, t_{F S}, t_{G S}, t_{F G S}$ (the letters in the subscript indicate the positive subformulas of the type). $\mathcal{E}_{\varphi}$ contains:

$$
\begin{aligned}
x_{\emptyset}+x_{F}+x_{G}+x_{S}+x_{F G}+x_{F S}+x_{G S}+x_{F G S} & =122 \\
x_{F}+x_{F G}+x_{F S}+x_{F G S} & =50 \\
x_{\emptyset} & =0
\end{aligned}
$$

## Systems of inequalities - Example

$\varphi:=\exists={ }^{122} x$ ヤ $\uparrow \wedge \exists^{=50} x$ French $(x) \wedge \exists^{=36} x \operatorname{German}(x) \wedge \exists^{=36} x \operatorname{Spanish}(x) \wedge$
$2^{(\varphi)}$

$$
\forall x(\operatorname{French}(x) \vee \operatorname{German}(x) \vee \operatorname{Spanish}(x)) \wedge
$$

$$
\exists=38 x(\operatorname{French}(x) \wedge \neg \operatorname{German}(x)) \wedge
$$

$$
|\varphi|^{\prime} \exists=10 x(\operatorname{French}(x) \wedge \operatorname{German}(x) \wedge \operatorname{Spanish}(x))
$$

Denote the 1-types over the signature French, German, Spanish by $t_{\emptyset}, t_{F}, t_{G}, t_{S}, t_{F G}, t_{F S}, t_{G S}, t_{F G S}$ (the letters in the subscript indicate the positive subformulas of the type). $\mathcal{E}_{\varphi}$ contains:

$$
\left\{\begin{aligned}
x_{\emptyset}+x_{F}+x_{G}+x_{S}+x_{F G}+x_{F S}+x_{G S}+x_{F G S} & =122 \\
x_{F}+x_{F G}+x_{F S}+x_{F G S} & =50 \\
x_{\emptyset} & =0 \\
x_{F}+x_{F S} & =38
\end{aligned}\right.
$$

## $\mathcal{E}_{\varphi}$ FOR OUR EXAMPLE

$$
\begin{aligned}
x_{\emptyset}+x_{F}+x_{G}+x_{S}+x_{F G}+x_{F S}+x_{G S}+x_{F G S} & =122 \\
x_{F}+x_{F G}+x_{F S}+x_{F G S} & =50 \\
x_{\emptyset} & =0 \\
x_{F}+x_{F S} & =38 \\
x_{G}+x_{F G}+x_{G S}+x_{F G S} & =36 \\
x_{S}+x_{F S}+x_{G S}+x_{F G S} & =36 \\
x_{F S}+x_{F G S} & =18 \\
x_{G S}+x_{F G S} & =21 \\
x_{F G S} & =10
\end{aligned}
$$

Lemma: $\mathcal{E}_{\varphi}$ has a non-negative integer solution iff $\varphi$ has a model.

## SYSTEMS OF INEQUALITIES - FORMALIZED

$$
\varphi:=\bigwedge_{i=1}^{m} \exists \bowtie C_{i} x \theta_{i}
$$

Let $\sigma=\left\{P_{1}, \ldots, P_{l}\right\}$. A 1-type (over $\sigma$ ) is any of the formulas:

$$
\pm P_{1} x \wedge \ldots \wedge \pm P_{l} x
$$

Let $\mathfrak{A}$ be a finite $\sigma$-structure and $t_{1}, \ldots, t_{L}$ be an enumeration of all 1-types, $L=2^{l}$. We characterize $\mathfrak{A}$ by the sequence of natural numbers $\left(\alpha_{1}, \ldots, \alpha_{L}\right)$ where $a_{j}=\left|\left\{a \in A: \mathfrak{A} \vDash t_{\boldsymbol{j}}(a)\right\}\right|$. The system $\mathcal{E}_{\varphi}$ contains for each conjunct $\exists \bowtie_{i} C_{i} x \theta_{i}$ the inequality:

$$
c_{i, 1} x_{1}+\ldots+c_{i, L} x_{L} \bowtie_{i} C_{i},
$$

where $c_{i, j}=1$ if the 1-type $t_{j}$ entails $\theta_{i}$ and $c_{i, j}=0$, otherwise.

## COMPLEXITY OF $\mathcal{C}^{1}$

## Lemma (Reduction property)

$\mathcal{E}_{\varphi}$ has a non-negative integer solution iff $\varphi$ has a model. Moreover, every solution of $\mathcal{E}_{\varphi}$ characterizes some model of $\varphi$.

The problem integer programming is as follows:

- given: a system $\mathcal{E}$ of linear equations and inequalities check whether $\mathcal{E}$ has a solution over $\mathbb{N}$.

Theorem (Borosh and Treybig 1976)
Integer programming is in NPTIME.
$\mathcal{E}_{\varphi}$ has $m$ inequalities and $L=2^{l}$ variables. Recall $m, l \leq|\varphi|$.


$$
\varphi:=\bigwedge_{i=1}^{m} \exists \bowtie c_{i x} x \theta_{i}
$$

## Optimal Complexity for $\mathcal{C}^{1}$

$\varphi:=\bigwedge_{i=1}^{m} \exists \bowtie C_{i} x \theta_{i} ; \quad \mathcal{E}_{\varphi}: m$ inequalities, $L=2^{l}$ variables.
Lemma (linear algebra)
If $\mathcal{E}_{\varphi}$ has a solution over $\mathbb{N}$, then $\mathcal{E}_{\varphi}$ has a solution over $\mathbb{N}$ with at most $m \log (L+1)$ non-zero entries.
Corollary
$S A T\left(\mathcal{C}^{1}\right) \in N P$.
polynomial in $|\varphi|$
Proof.
Let $C=\max \left\{C_{i}: 1 \leq i \leq m\right\}$. If $\left(\alpha_{1}, \ldots, \alpha_{L}\right)$ is a solution of $\mathcal{E}_{\varphi}$, then so is $\left(\beta_{1}, \ldots, \beta_{L}\right)$, where $\beta_{j}=\min \left(\alpha_{j}, C\right)$.
The linear algebra Lemma allows one to first guess a polynomial number of non-zero variables and write down the system $\mathcal{E}_{\varphi}$ only over these variables; since Integer Programming is in NPTIME, solutions of such systems can be guessed and verified in time bounded by a polynomial function of $|\varphi|$.

## Reduction to Integer Programming

## $C^{2} J_{y}^{=1} R(x, y)$

Idea: Depending on the logic:
 identify (finitely many types of) Wuilding blocks of a potential model and connecting conditions for them, describe them in a succinct way by a set of (in)equalities.
Advantages:

- Useful for solving simultaneously SAT and FINSAT. We look for solutions over $\mathbb{N}$ (FINSAT) or over $\mathbb{N} \cup\{\infty\}$ (SAT), e.g.

$$
x+1=x
$$

has a solution $x=\infty$.

- Gives better (optimal) complexity bounds.

We will see more about this approach later in the course.
$\mathcal{I} O^{2}$ : Two -variable fragment of $F O$
$x, y$
relational symbols of arity $\leq 2$ no constants

Def. Normal forms for $\mathrm{FD}^{2}$

$$
\begin{aligned}
& \forall y(S(y) \leftrightarrow \exists x(A(x) \wedge B(y))) \equiv \\
& \forall y[(S(y) \rightarrow \exists x A(x) \wedge B(y))] \wedge[(\exists x A(x) \wedge B(y)) \rightarrow S(y)] \equiv \\
& =\left(\forall x \exists_{y} \delta(x) \rightarrow A(y) \wedge B(x)\right) \wedge\left(\forall y\left(\exists_{x} A(x) \wedge B(y)\right) \rightarrow S(y)\right) \\
& \forall y \quad \sim(-11-) \vee S(y) \\
& \forall y\left(\begin{array}{cc}
\forall x & 1 A(x) \cup B(y) \\
11
\end{array}\right) \vee S(y) \\
& \forall x \forall y \quad(\neg A(x) \vee \neg B(y) \vee S(y))
\end{aligned}
$$

