Introduction into Mathematics of Constraint Satisfaction

Andrei Krokhin
Durham University

Part I

Two Well-Known Problems

SAT: is a given propositional formula in CNF satisfiable?

$$F = (\neg x \vee y \vee \neg z) \wedge (x \vee y \vee \neg z) \wedge (\neg x \vee \neg y \vee z)$$

Linear Equations: does a given system of linear equations have a solution in the fixed field K?

$$\begin{cases} 2x + 2y + 3z = 1 \\ 3x - 2y - 2z = 0 \\ 5x - y + 10z = 2 \end{cases}$$

The Constraint Satisfaction Problem

CSP

Instance: (V, D, C) where

- V is a finite set of variables
- D is a set of values (aka domain)
- C is a finite set of constraints $\{C_1, \ldots, C_q\}$
  - each constraint C_i is a pair $(\overline{s}_i, R_i)$ with
    - scope $\overline{s}_i$ - a list of variables of length m_i,
    - relation R_i - an m_i-ary relation over D

Question: Is there $f : V \to D$ s.t. $f(\overline{s}_i) \in R_i$ for all i?
Can ask to find such an f too.

Some Real-World Examples of CSPs

- Solving a Sudoku puzzle
- Drawing up timetable for a conference
- Choosing frequencies for a mobile-phone network
- Fitting a protein structure to measurements
- Laying out components on a circuit board
- Finding a DNA sequence from a set of contigs
- Scheduling a construction project

Outline of the Course

1. The CSP and its forms
   - Examples of CSPs from
     - Logic
     - Algebra
     - Graph Theory
     - AI
   - Complexity Issues
2. Computational questions
   - What questions do we ask about CSPs?
3. Mathematical techniques
   - What maths do we use to analyse those questions?

Relational Structures

A signature τ is

- a finite sequence $(R_1, \ldots, R_k)$ of relation symbols
- each R_i has an associated arity $ar(R_i)$.

A relational structure of signature τ (or τ-structure) is

- a tuple $\mathcal{A} = (A; R_1^{\mathcal{A}}, \ldots, R_k^{\mathcal{A}})$ where
  - A is a set called the universe of $\mathcal{A}$
  - each $R_i^{\mathcal{A}}$ is a relation on A of arity $ar(R_i)$

If $\tau = \{E\}$ and $\mathrm{ar}(E) = 2$ then τ-structures are digraphs.

CSP in Logical Setting

CSP

Instance: A τ-structure $\mathcal{B}$ and a formula $\exists x_1 \ldots \exists x_n \varphi$

where $\varphi(x_1, \ldots, x_n) = R_{i_1}(\overline{s}_1) \wedge \ldots \wedge R_{i_q}(\overline{s}_q)$.

Question: Does $\mathcal{B}$ satisfy φ?

The $\overline{s}_i$'s = constraint scopes $\overline{s}_i$

Predicates $R_i^{\mathcal{B}}$ = constraint relations R_i

In Database Theory, Conjunctive-Query Evaluation

Exercise

Let $\mathcal{B} = (\{0, 1, 2\}; E)$ where $E = \{(0, 1), (1, 2), (2, 0)\}$.

Let $\varphi_1 = E(x_0, x_1) \wedge E(x_1, x_2) \wedge E(x_2, x_3) \wedge E(x_3, x_0)$.

Let $\varphi_2 = E(x_0, x_1) \wedge E(x_1, x_2) \wedge E(x_2, x_3) \wedge E(x_3, x_1)$.

Does $\mathcal{B}$ satisfy $\exists x_0 \ldots \exists x_3 \varphi_1$? Does it satisfy $\exists x_0 \ldots \exists x_3 \varphi_2$?

CSP in Combinatorial Setting

CSP

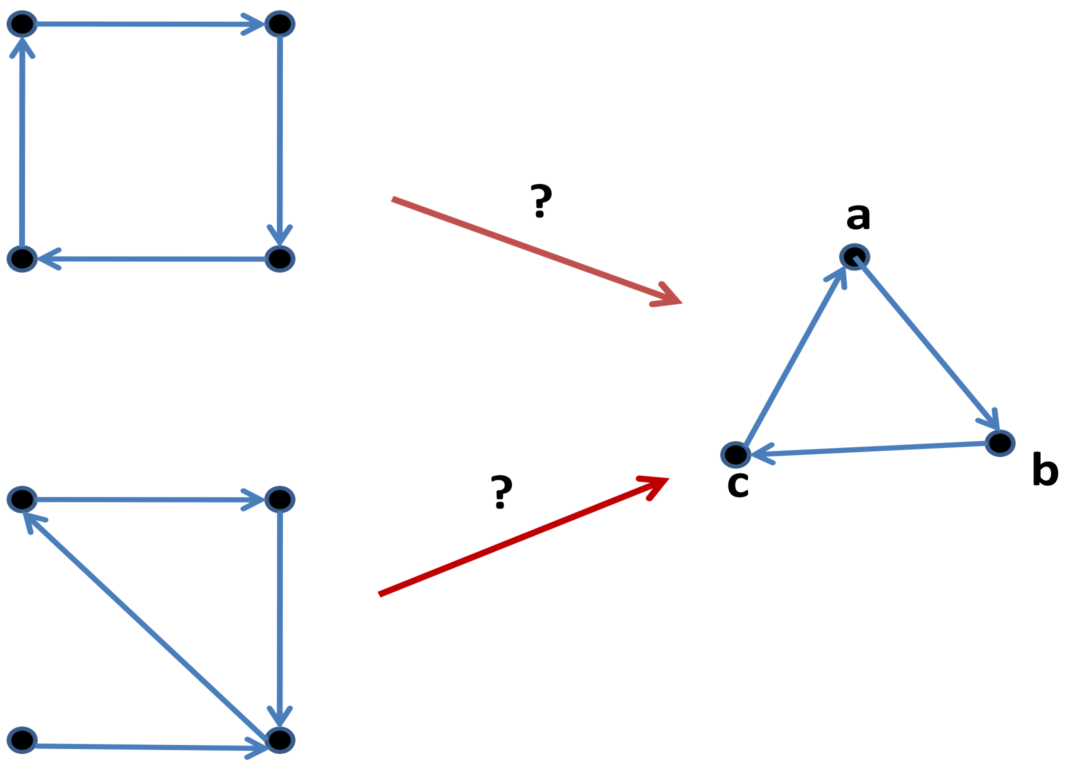
Instance: Two τ-structures, $\mathcal{A}$ and $\mathcal{B}$

Question: Is there a homomorphism $h : \mathcal{A} \to \mathcal{B}$?

$$\forall i \ [(a_1, \ldots, a_{n_i}) \in R_i^{\mathcal{A}} \Longrightarrow (h(a_1), \ldots, h(a_{n_i})) \in R_i^{\mathcal{B}}]$$

- Think of elements in $\mathcal{A}$ as of variables.
  Tuples in relations in $\mathcal{A}$ = constraint scopes $\overline{s}_i$.
- Think of elements in $\mathcal{B}$ are values.
  Relations in $\mathcal{B}$ = constraint relations R_i.

Exercise



Example: 2-SAT in Hom Form

Let $R^{\mathcal{B}}_{ab} = \{0,1\}^2 \setminus \{(a,b)\}$ and $\mathcal{B} = (\{0,1\}; R^{\mathcal{B}}_{00}, R^{\mathcal{B}}_{01}, R^{\mathcal{B}}_{11})$.

An instance of 2-SAT, say

$$F = (\neg x \vee \neg z) \wedge (x \vee y) \wedge (y \vee \neg z) \wedge (u \vee x) \wedge (x \vee \neg u) \dots$$

becomes a structure $\mathcal{A}$ with base set $\{x, y, z, u, \dots\}$ and

$R^{\mathcal{A}}_{00} = \{(x,y), (u,x), \dots\}$

$R^{\mathcal{A}}_{01} = \{(y,z), (x,u), \dots\}$

$R^{\mathcal{A}}_{11} = \{(x,z), \dots\}$

Then $h : \mathcal{A} \to \mathcal{B}$ iff h is a satisfying assignment for F.

Recap: 3 Forms of CSP

- Variable-value (AI, Algebra)

  Given finite sets A (variables), B (values), and a set of constraints $\{(\overline{s}_1, R_1), \ldots, (\overline{s}_q, R_q)\}$ over A, is there a function $\varphi : A \to B$ such that $\varphi(\overline{s}_i) \in R_i$ for all i?

- Satisfiability (Logic, Databases)

  Given a finite structure (or database) $\mathcal{B}$ and a $\exists\wedge$-FO sentence (or conjunctive query) φ, does $\mathcal{B}$ satisfy φ?

- Homomorphism (Logic, Graph Theory)

  Given two similar relational structures,
  $\mathcal{A} = (A;\ R_1^{\mathcal{A}}, \ldots, R_k^{\mathcal{A}})$ and $\mathcal{B} = (B;\ R_1^{\mathcal{B}}, \ldots, R_k^{\mathcal{B}})$,
  is there a homomorphism $h : \mathcal{A} \to \mathcal{B}$?

Representation Issues

When speaking about algorithms/complexity

- for finite domains, will usually assume that relations are given explicitly, by full list of tuples;
- for infinite domains, need to
  - fix the domain in advance,
  - give relations in some finite form.

Propositional Satisfiability

Important fragments of SAT:

- 2-SAT — clauses with at most 2 var's tractable
- HORN k-SAT — clauses $(\neg x_1 \vee \ldots \vee \neg x_n)$, $(\neg x_1 \vee \ldots \vee \neg x_n \vee x_{n+1})$, and (x) tractable
- 1-IN-3-SAT — one constraint type $\{(1,0,0),(0,1,0),(0,0,1)\}$ **NP**-complete
- NOT-ALL-EQUAL-SAT — one constraint type $\{0,1\}^3 \setminus \{(0,0,0),(1,1,1)\}$ **NP**-complete

Equations over Groups

A group is an algebra $(G; \cdot, ^{-1}, 1)$ such that $1 \cdot x = x \cdot 1$, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, and $x \cdot x^{-1} = x^{-1} \cdot x = 1$

System of equations over a finite group G (Eq^*_G):

$$\begin{cases} a_1 x_{11} \cdots a_m x_{1m} a_{m+1} = 1 \\ \ldots \\ b_1 x_{n1} \cdots b_m x_{nm} b_{m+1} = 1 \end{cases}$$

Can replace $xyz\ldots = 1$ by $xy = x'$ and $x'z\ldots = 1$.

Theorem 1 (Goldmann, Russell '2002)
The problem Eq^*_G is tractable if G is Abelian and **NP**-*complete otherwise.*

Equations over Semigroups

A semigroup is an algebra $(G; \cdot)$ with $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

System of equations over a finite semigroup S (Eq^*_S):

can assume all equations of the form $u \cdot v = w$ where each of u, v, w is either a constant or a variable.

Theorem 2 (Klima,Tesson,Thérien'2005)
Assume that S is a monoid (has 1). If S is commutative and is the union of its subgroups then the problem Eq^*_S is tractable. Otherwise it is **NP***-complete.*

No full answer for general semigroups yet ...

Equations over Algebras

A finite algebra is a tuple $\mathbf{A} = (A; f_1, \ldots, f_k)$ where each f_i is an operation $f_i : A^{n_i} \to A$.

Can consider systems of equations over $\mathbf{A}$.
By the same trick, assume all equations are of the form $f_i(u_1, \ldots, u_{n_i}) = w$ where each of $u_1, \ldots, u_{n_i}, w$ is either a constant or variable.

Interesting work by Larose,Zádori'06 and Zádori'07+'11 – classification for large classes of algebras.

Colourability

k-COLOURABILITY

Instance: A graph $G = (V, G)$ and a number k.

Question: Is there a colouring of V in k colours such that adjacent vertices are different colour?

- Equivalent to CSP instance (G, K_k), in hom form. Indeed a required colouring is a homomorphism $G \to K_k$
- tractable for $k \leq 2$, **NP**-complete for $k \geq 3$.

List Colourability

List k-Colourability

Instance: A graph $G = (V, G)$, a number k, and a list L_v of allowed colours for each $v \in V$.

Question: Is there a k-colouring of G such that each vertex gets an allowed colour?

- Equivalent to CSP instance $(\mathcal{A}, \mathcal{B})$, in hom form
  - $\mathcal{B} = (\{1, \ldots, k\}; \neq_k, L_1, \ldots, L_{2^k})$ where $L_1, \ldots, L_{2^k}$ is a fixed enumeration of subsets of $\{1, \ldots, k\}$.
  - $\mathcal{A} = (V; E, U_1, \ldots, U_{2^k})$ where each U_i consists of vertices whose list is L_i.

Clique

CLIQUE

Instance: A graph G and a number k.

Question: Does G contain a k-clique, i.e. k pairwise adjacent vertices?

- Equivalent to CSP instance (K_k, G) in hom form. Indeed, G has a k-clique iff $K_k \to G$.
- **NP**-complete

Hamiltonian Circuit

HAMILTONIAN CIRCUIT

Instance: A graph $G = (V, G)$.

Question: Is there a cyclic ordering of V such that every pair of successive nodes in the ordering are adjacent in G?

- Equivalent to CSP instance $(\mathcal{A}, \mathcal{B})$ in hom form where
  - $\mathcal{A} = (V; C_V, \neq_V)$, C_V is a cyclic permutation on V,
  - $\mathcal{B} = (V; E, \neq_V)$
- **NP**-complete

Graph Isomorphism

GRAPH ISOMORPHISM

Instance: Two graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$.

Question: Is there an isomorphism (bijective homomorphism) from G_1 to G_2?

- Equivalent to CSP instance $(\mathcal{A}, \mathcal{B})$ in hom form where
  - $\mathcal{A} = (V; E_1, \overline{E_1})$ and $\mathcal{B} = (V; E_2, \overline{E_2})$
  - here $\overline{E_i} = \{(u, v) \in V^2 \mid (u, v) \notin E_i \text{ and } u \neq v\}$
- not known to be tractable or **NP**-complete
- main candidate for **NP**-intermediate

Directed *st*-Reachability

> DIRECTED *st*-REACHABILITY
>
> **Instance:** A digraph $G = (V, E)$ and two nodes s, t in it.
>
> **Question:** Is there a directed path from s to t in G?

- Essentially equivalent to the complement of $(\mathcal{A}, \mathcal{B})$
- $\mathcal{A} = (V; E, \{s\}, \{t\})$ and $\mathcal{B} = (\{0, 1\}, \leq, \{1\}, \{0\})$.
- tractable, **NL**-complete.

Graph Factors

Let G be a graph.

- A factor of G is a subgraph obtained by deleting edges.
- If X is a set of numbers, an X-factor of G is a factor such that the degree of each vertex belongs to X.
- What are 1-factors of a graph?
- What are 2-factors? $\{1, 2\}$-factors?
- Large subarea of graph theory.
- Recent result: if r is odd and k even with $2 \leq k < r/2$ then each r-regular graph has a $\{k, r - k\}$-factor.
- Open: does every 5-regular graph have a $\{1, 4\}$-factor?

Factors in Regular Graphs

Let $G = (V, E)$ be r-regular and let $X \subseteq \{0, 1, \ldots, r\}$.

Let $R_X = \{\mathbf{a} \in \{0, 1\}^r \mid weight(\mathbf{a}) \in X\}$.

Consider the following CSP instance $I = (V', D', C')$:

- $V' = E$, $D' = \{0, 1\}$, and $C' = \{C_v \mid v \in V\}$;
- each $C_v = ((e_1, \ldots, e_r), R_X)$ where $e_1, \ldots, e_r$ are the edges incident to v.

Fact. G has an X-factor iff I has a solution.

Fact. There is a 1-1 correspondence between X-factors in r-regular graphs and solutions of Boolean CSP instances that use only relation R_X and each variable appears twice.

Vertex Cover

VERTEX COVER

Instance: A graph $G = (V, E)$ and a number k.

Question: Is there a subset $V' \subseteq V$ such that $|V'| \leq k$ and, for each $(a, b) \in E$, we have $a \in V'$ or $b \in V'$?

- Equivalent to CSP instance $(V, \{0, 1\}, C)$,
- $C = \{u \vee v \mid (u, v) \in E\}$,
- additionally: want only solutions with k ones.
- **NP**-complete

Independent Set

INDEPENDENT SET

Instance: A graph $G = (V, E)$ and a number k.

Question: Is there a subset $V' \subseteq V$ such that $|V'| \geq k$ and $(V')^2 \cap E = \emptyset$?

- Equivalent to CSP instance $(V, \{0, 1\}, C)$,
- $C = \{\neg u \vee \neg v \mid (u, v) \in E\}$,
- additionally: want only solutions with k ones.
- **NP**-complete

Parameterized Complexity in One Slide

- Both VERTEX COVER and INDEPENDENT SET can be solved in $O(n^k)$ by exhaustive search.
- VERTEX COVER can be solved in $O(2^k \cdot n^2)$ — good
- $O(f(k) \cdot n^c)$ algorithm — fixed-parameter tractable (FPT)
- $\mathbf{W}[\mathbf{1}]$ — parameterized analog of **NP** (kind of)
- IND SET is $\mathbf{W}[\mathbf{1}]$-complete — $n^{o(k)}$ algorithm unlikely.

Biclique

BICLIQUE

Instance: A graph $G = (V, E)$ and a number k.

Question: Does G contain an induced subgraph isomorphic to $K_{k,k}$ (bipartite k-clique)?

- **NP**-complete. Open problem: Is it FPT?
- Equivalent to CSP instance $(V, \{0, 1, 2\}, C)$
- where $C = \{C'_{(u,v)} \mid (u, v) \in E\} \cup \{C''_{(u,v)} \mid (u, v) \notin E\}$,
- $C'_{(u,v)} = ((u, v), R')$, $R = \{0, 1, 2\}^2 \setminus \{(0, 0), (1, 1)\}$,
- $C''_{(u,v)} = ((u, v), R'')$, $R = \{0, 1, 2\}^2 \setminus \{(0, 1), (1, 0)\}$,
- want only solutions with k zeroes and k ones.

Directed Acyclicity

ACYCLIC DIGRAPH

Instance: A digraph $G = (V, E)$

Question: Is it true that G has no direct cycles?

- Equivalent to CSP instance $(V, \mathbb{Q}, C)$ where
- $C = \{C_e \mid e \in E\}$, $C_{(u,v)} = ((u, v), <)$.
- tractable

Betweenness

BETWEENNESS

Instance: A finite set V and a set $M \subseteq V^3$.

Question: Is there a linear ordering of V such that, for each triple $(u, v, w) \in M$, we have $u < v < w$ or $u > v > w$?

- Equivalent to CSP instance $(V, \mathbb{Q}, C)$ where
- $C = \{C_m \mid m \in M\}$, $C_m = ((u, v, w), R_b)$, $R_b = \{(a, b, c) \in \mathbb{Q}^3 \mid a < b < c \text{ or } a > b > c\}$.
- **NP**-complete

Ordering CSP

Let Π be a set of permutations on $\{1, \ldots, k\}$.

Π-Ordering CSP

Instance: A finite set V and a family $M \subseteq V^k$.

Question: Is there a linear order on V such that each element of M agrees with some permutation from Π?

- Straightforward generalisation of the previous two.
- Directed Acyclicity — $k = 2$, $\Pi = \{(12)\}$
- Betwenness — $k = 3$, $\Pi = \{(123), (321)\}$

Allen’s Interval Algebra

The most popular formalism in temporal reasoning (AI) [Allen, 1983].

Allows qualitative binary constraints between time intervals.

Allen's Algebra: 13 Basic Relations

x precedes y	p	xxx
y preceded by x	p^{-1}	yyy
x meets y	m	xxxx
y met by x	m^{-1}	yyyy
x overlaps y	o	xxxx
y overlapped by x	o^{-1}	yyyy
x during y	d	xxx
y includes x	d^{-1}	yyyyyyy
x starts y	s	xxx
y started by x	s^{-1}	yyyyyyy
x finishes y	f	xxx
y finished by x	f^{-1}	yyyyyyy
x equals y	$\equiv$	xxxx
		yyyy

AIA Satisfiability

Relations in AIA — 2^{13} disjunctions of the basic relations.

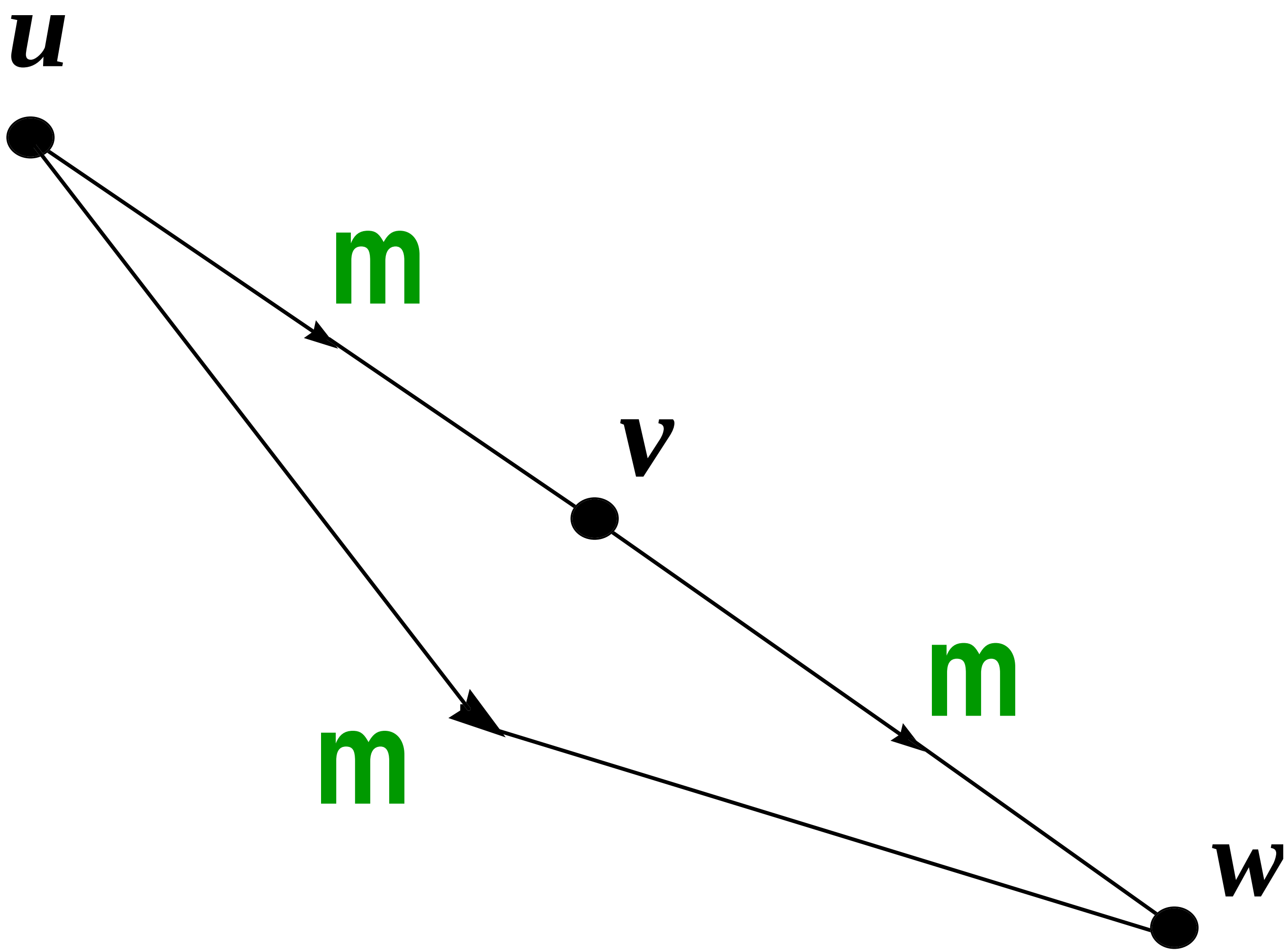
AIA-SAT

Instance: Given a labelled digraph $G = (V, A)$ where

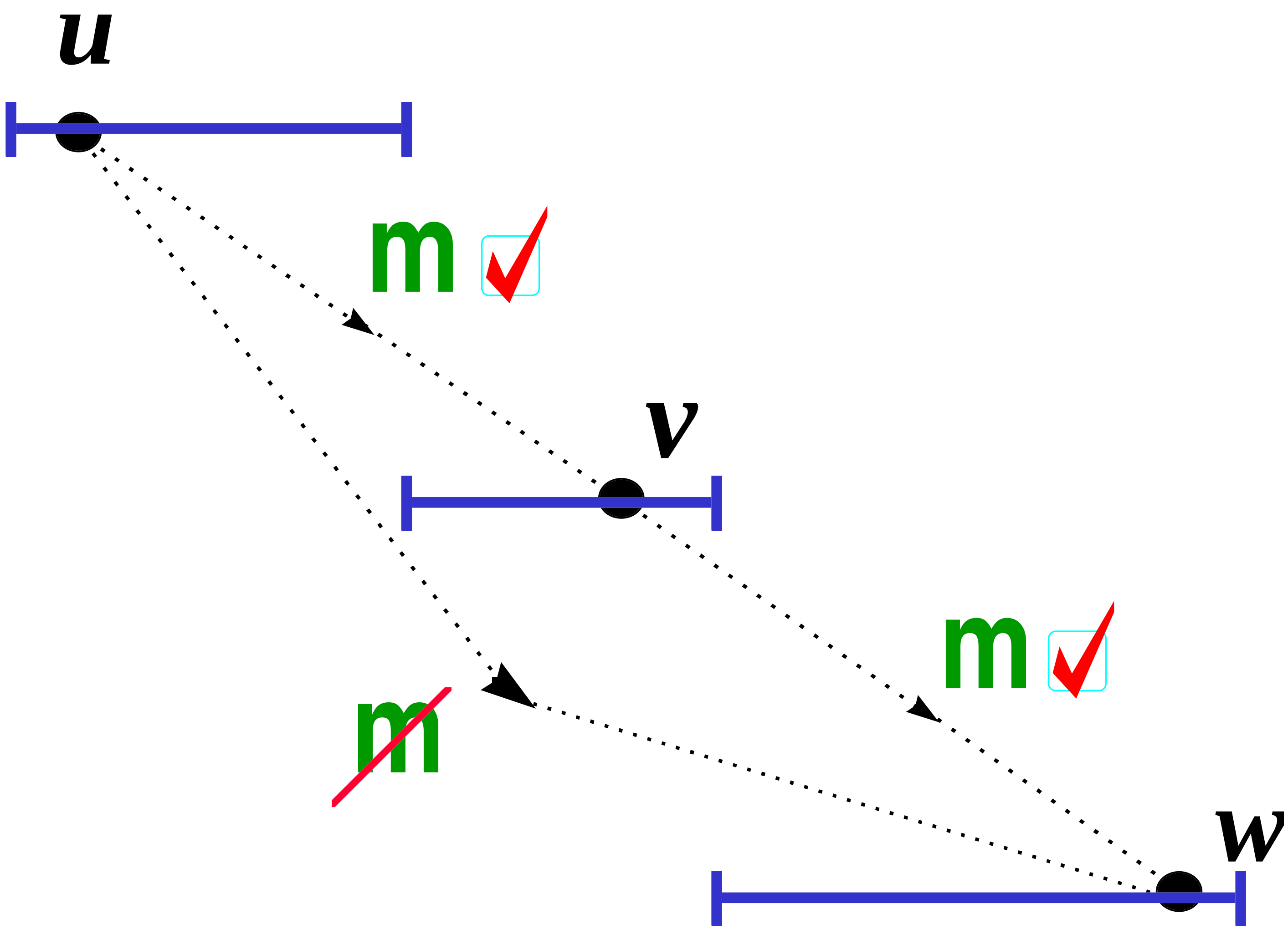
- V is a set of interval variables and
- A consists of triples (u, r, v) with $u, v \in V$ and r in AIA.

Question: Is there an assignment of intervals for the variables such that relations on all arcs from A are satisfied?

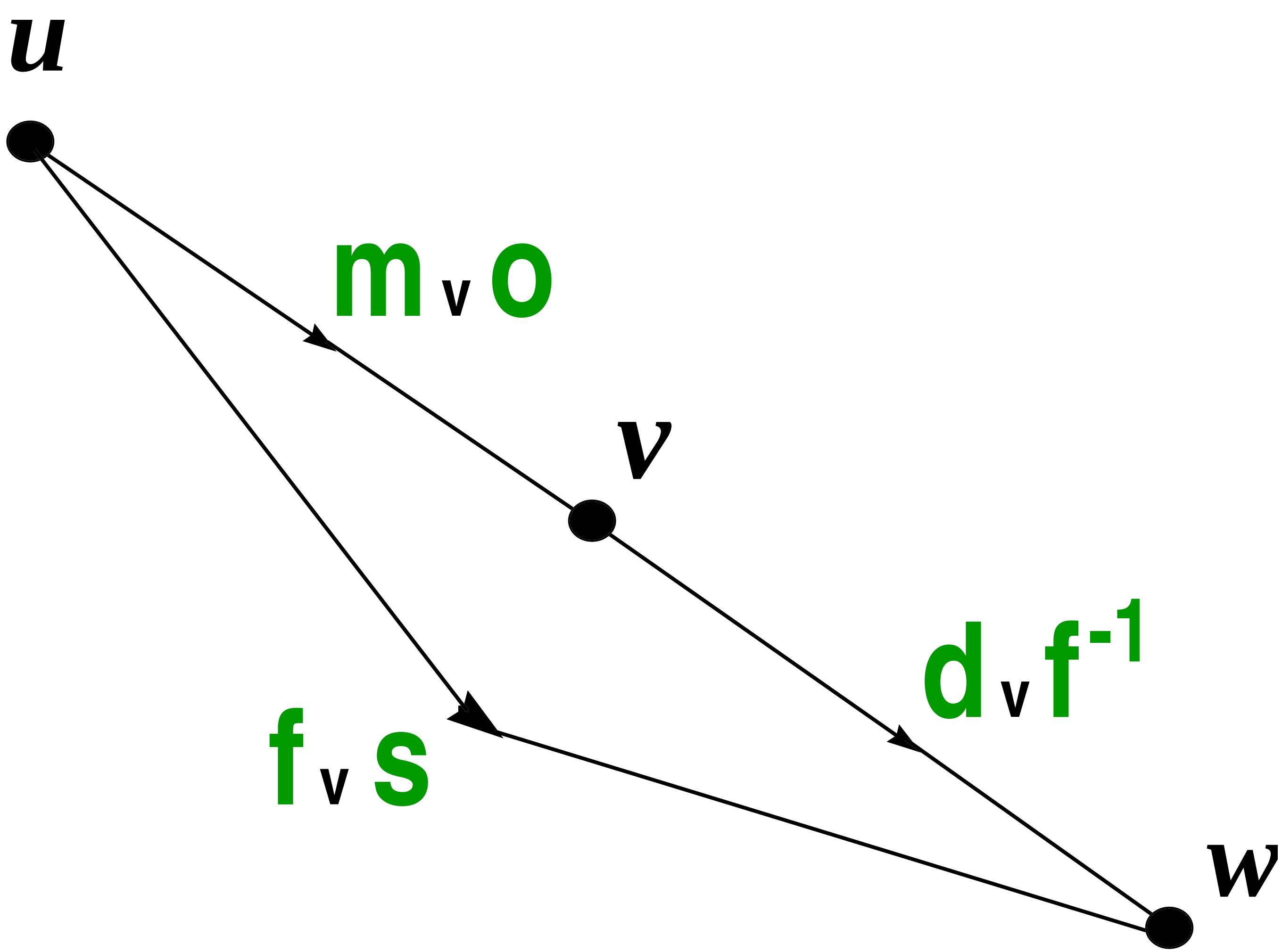
Example



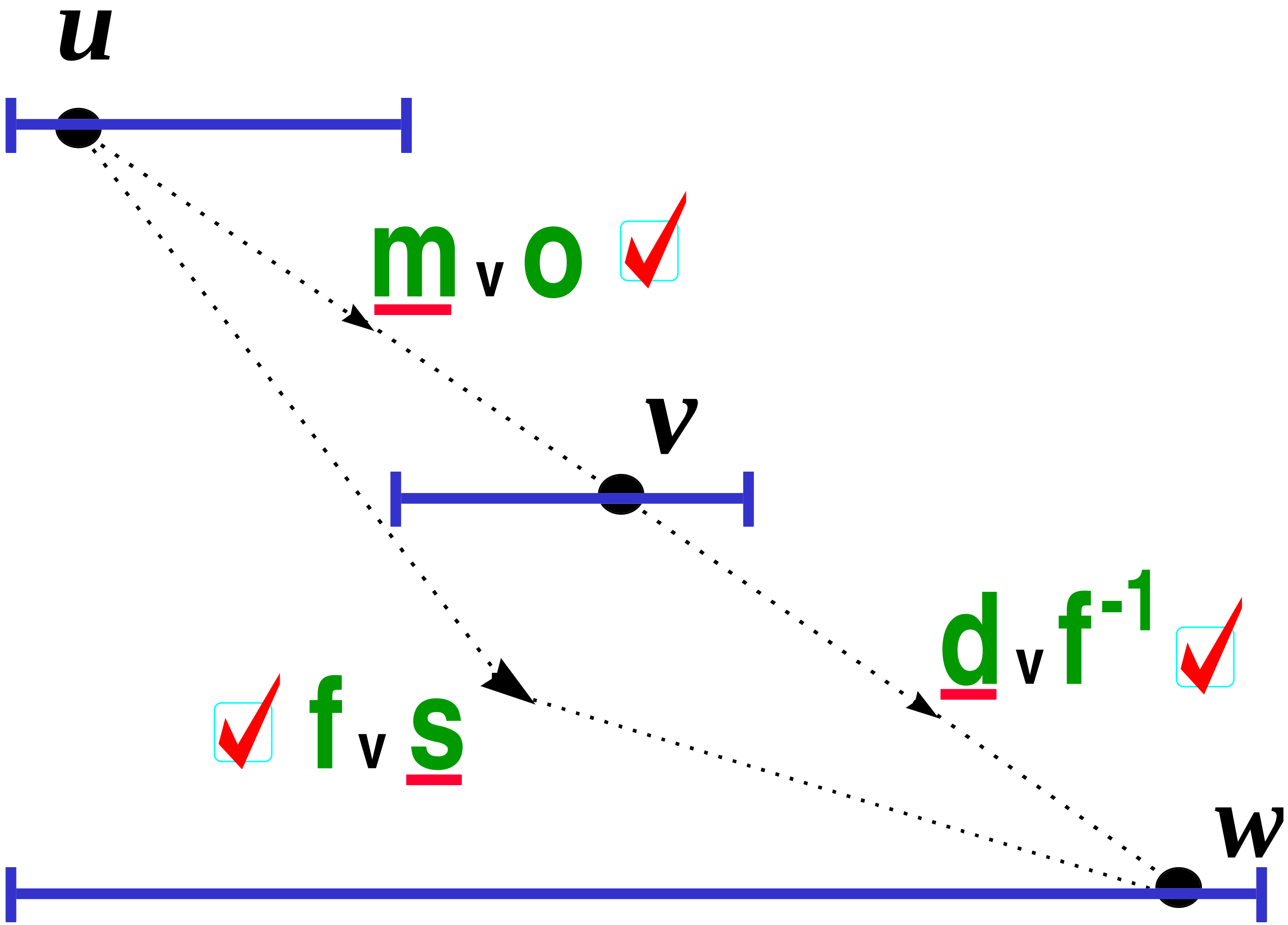
Example



Example



Example



Complexity Classifications

- Complexity Theory: aims at classifying combinatorial problems by computational complexity.
- Want: a test bed to see if they are good at this.
- CSP is very expressive and rich, but clean and manageable
- Natural test bed for complexity classifications (and algorithmic techniques)