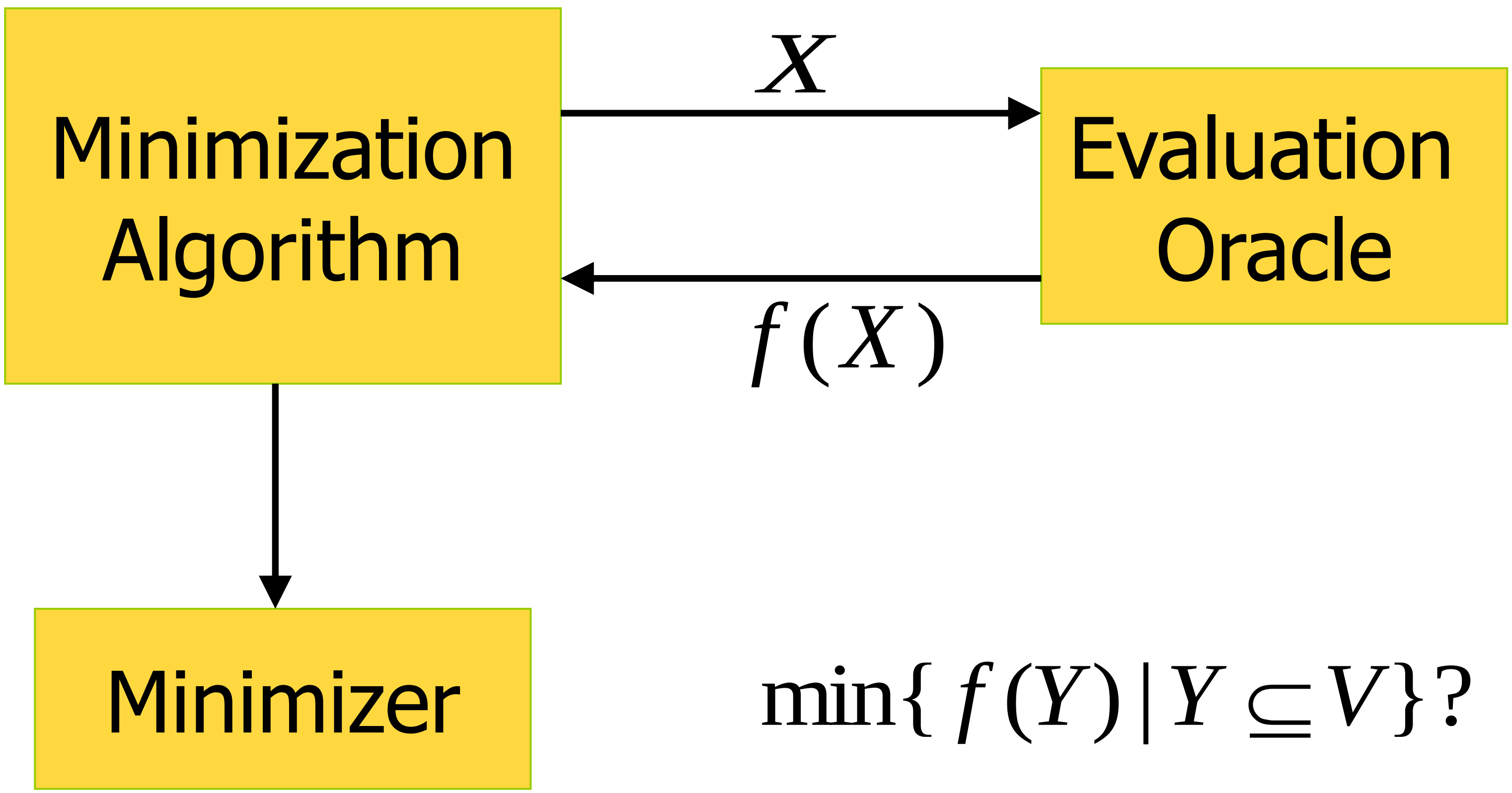
Submodular Function Minimization

Satoru Iwata

(University of Tokyo)

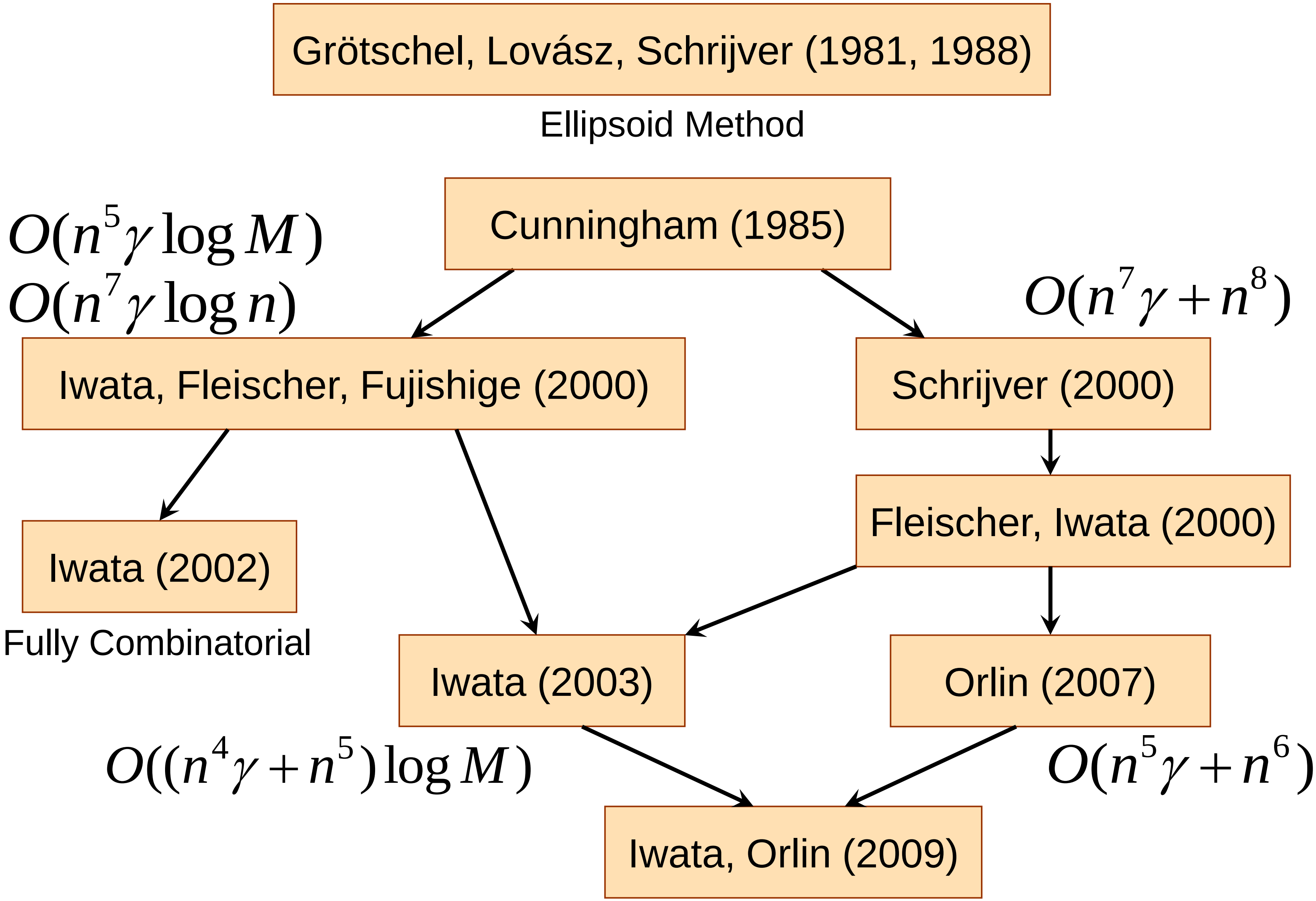
Submodular Function Minimization

Assumption: $f(\phi) = 0$



$$\min\{ f(Y) \mid Y \subseteq V \}?$$

Submodular Function Minimization



Submodular Functions and Polyhedra

$f : 2^V \to \mathbf{R}$ $\qquad\qquad$ $f(\phi) = 0$

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y), \quad \forall X, Y \subseteq V$$

Submodular Polyhedron

$$P(f) = \{x \mid x \in \mathbf{R}^V, \forall Y \subseteq V, x(Y) \leq f(Y)\}$$

Base Polyhedron

$$B(f) = \{x \mid x \in P(f), x(V) = f(V)\}$$

Min-Max Theorem

Theorem Edmonds (1970)

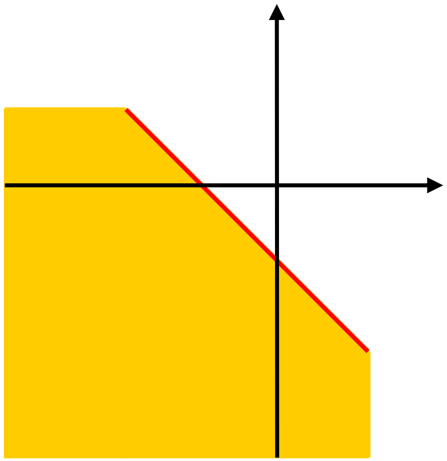
$$\min_{Y \subseteq V} f(Y) = \max\{ z(V) \mid z \in P(f), z \leq 0\}$$

$$= \max\{ x^{-}(V) \mid x \in B(f)\}$$

$$x^{-}(v) := \min\{0, x(v)\}$$

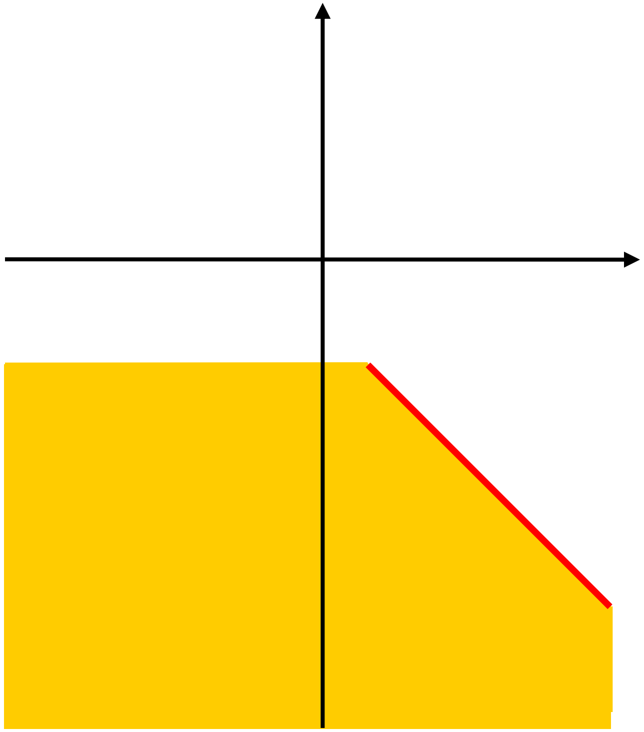
$$z(V) \leq z(Y) \leq f(Y)$$

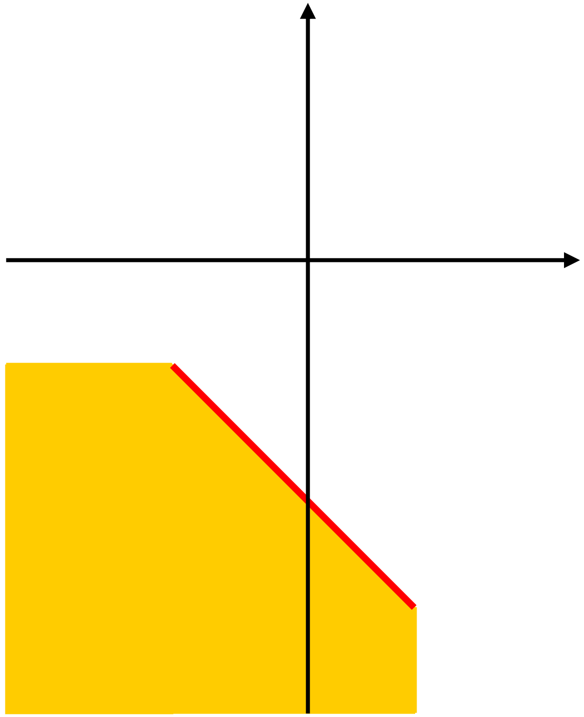
$$x^{-}(V) \leq x(Y) \leq f(Y)$$



Combinatorial Approach

$$\min_{Y \subseteq V} f(Y) = \max\{\, x^-(V) \mid x \in B(f)\}$$





Min-Max Theorem

$$\min_{Y \subseteq V} f(Y) = \max\{ z(V) \mid z \in P(f), z \leq 0\}$$

$$= \max\{ x^-(V) \mid x \in B(f)\}$$

$$f^\circ(X) := \min\{ f(Y) \mid Y \subseteq X\}$$ Submodular

$$P(f^\circ) \subseteq P(f)$$

$$z \in B(f^\circ) \Rightarrow z(V) = f^\circ(V) = \min_{Y \subseteq V} f(Y)$$

Combinatorial Approach

$$\boxed{\min_{Y \subseteq V} f(Y) = \max\{\, x^{-}(V) \mid x \in B(f)\}}$$

Convex Combination

$$x = \sum_{i \in I} \lambda_i y_i$$

$y_i \in B(f)$: Extreme Base Generated by the Greedy Algorithm with an Linear Ordering L_i in V.

IFF Scaling Algorithm

$$\delta \approx \frac{M}{n^2} \longrightarrow \delta < \frac{1}{n^2}$$

$\left\lfloor \dfrac{f_\delta(X)}{\delta} \right\rfloor$: Submodular

$$f_\delta(X) = f(X) + \underline{\delta |X| \cdot |V \setminus X|}$$

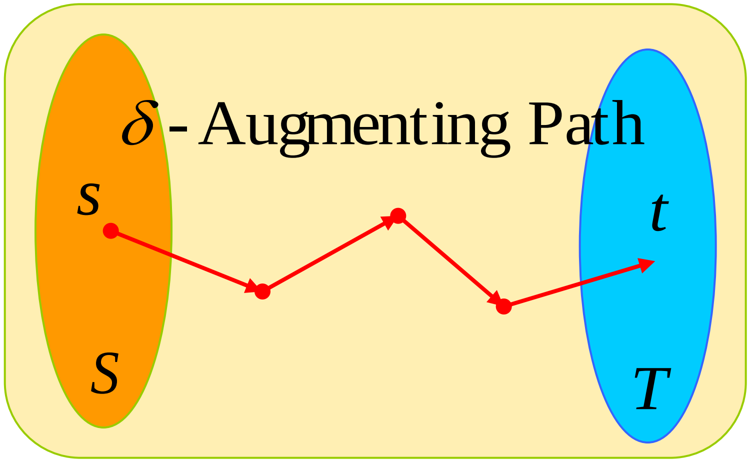
Cut Function

φ : Flow in the Complete Digraph $\varphi(u,v) \leq \delta$

$$x = \sum_{i \in I} \lambda_i y_i$$

$$\boxed{z = x + \partial \varphi}$$

Increase $z^-(V)$

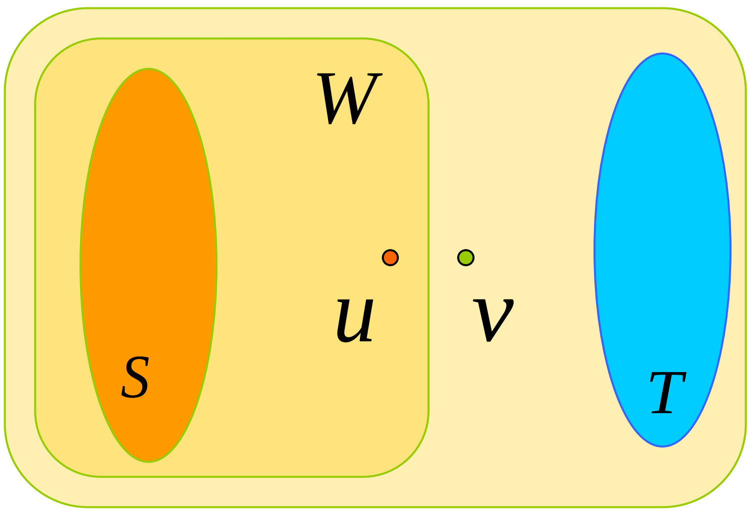


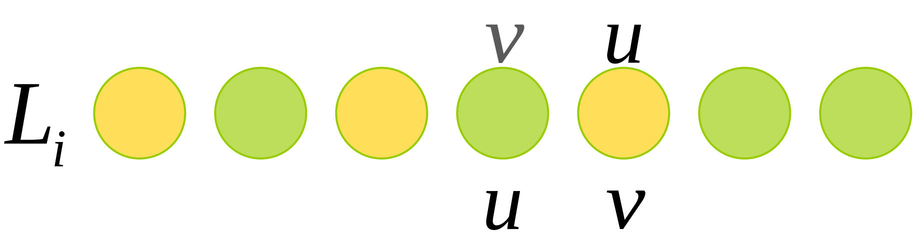
$S = \{v \mid z(v) < -\delta\}$ $T = \{v \mid z(v) > \delta\}$

IFF Scaling Algorithm

No Path from S to T

$$W := \{v \mid v : \text{Reachable from } S\}$$





Double-Exchange

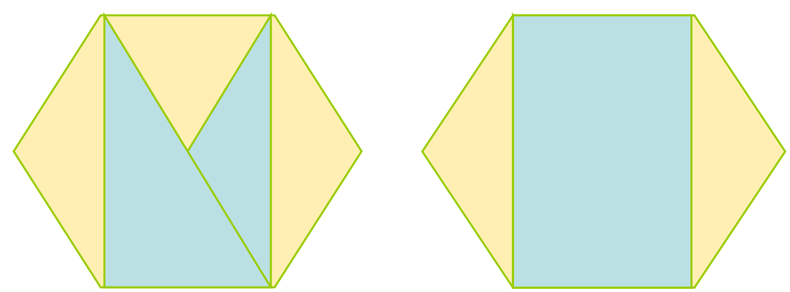
$$y_i := y_i + \beta(\chi_u - \chi_v)$$

$$\alpha := \min\{\delta, \lambda_i \beta\}$$

$$x := x + \alpha(\chi_u - \chi_v)$$

$$\varphi(u,v) := \varphi(u,v) - \alpha$$

$\alpha = \lambda_i \beta$ $\alpha < \lambda_i \beta$



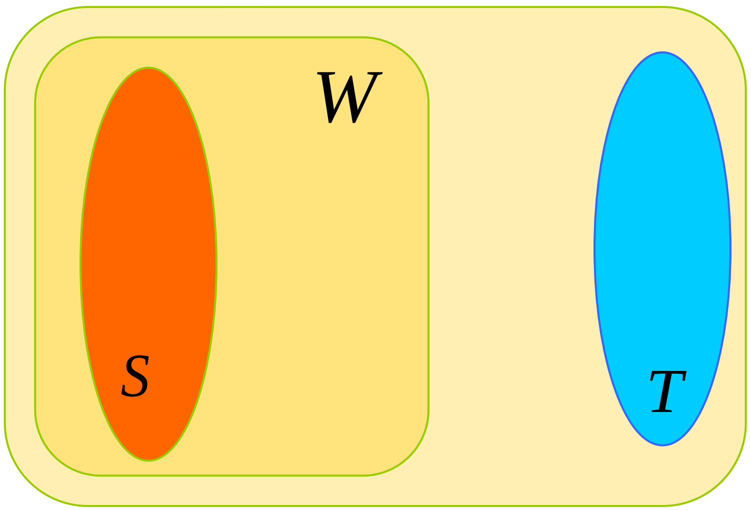
Saturating Nonsaturating

IFF Scaling Algorithm

No Path from S to T
No Active Triple(i,u,v)

$$y_i(W) = f(W), \ \forall i \in I$$

$$x(W) = \sum_{i \in I} \lambda_i y_i(W) = f(W)$$

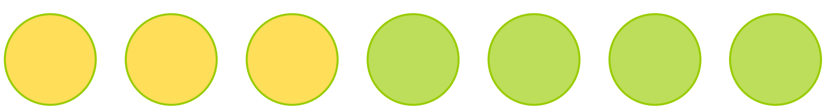


$$z^-(V) \geq f(W) - n\delta$$

$$x^-(V) \geq f(W) - n^2\delta$$

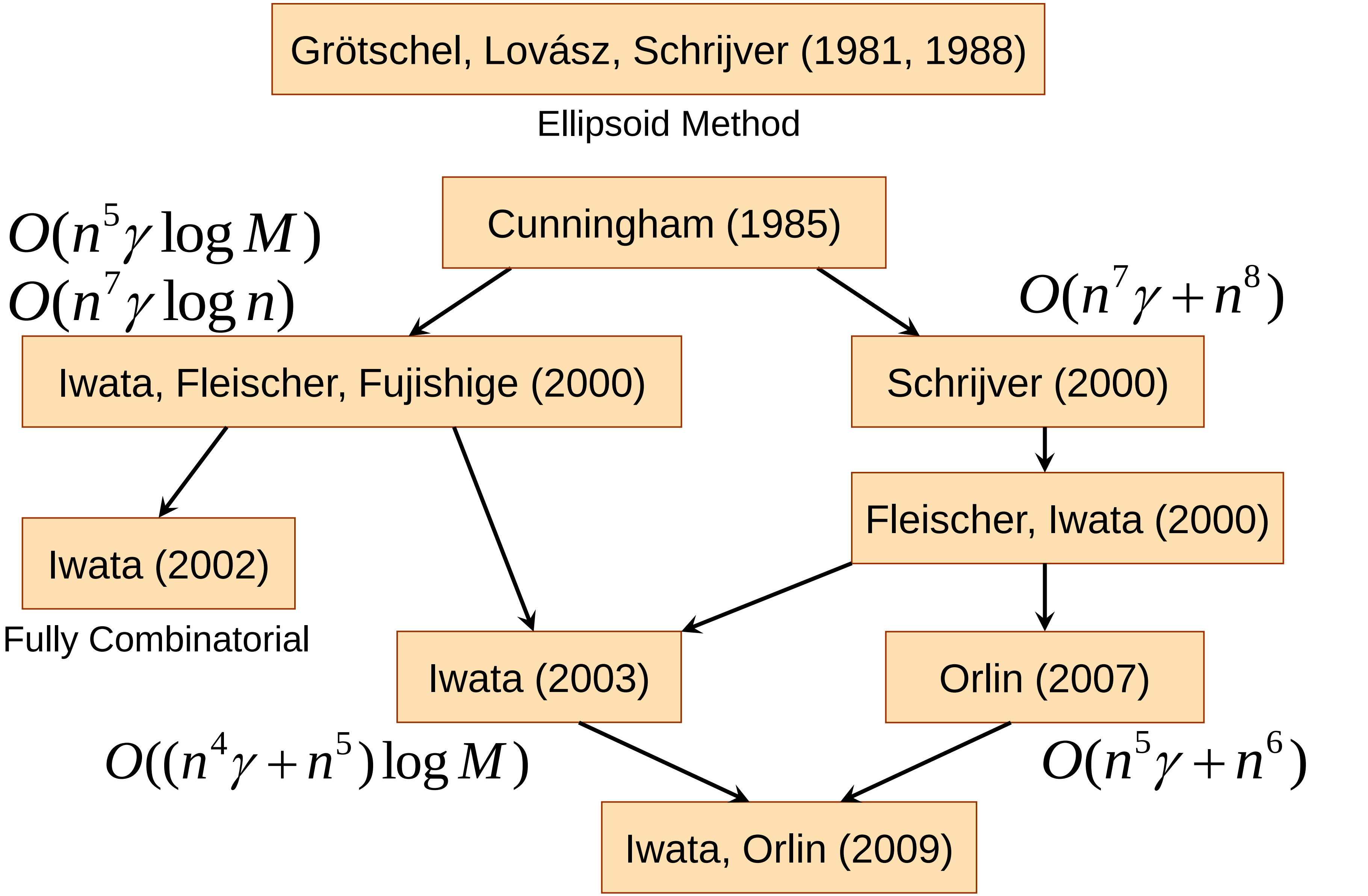
$\delta < \frac{1}{n^2} \Rightarrow f(W)$: Min

L_i



$$O(n^5 \gamma \log M)$$

Submodular Function Minimization



Fully Combinatorial Algorithm

Addition, Subtraction, Comparison

Oracle Call (Function Evaluation)

Multiplication $\alpha \in \mathbf{R}, \mu \in \mathbf{Z} \longrightarrow$ Compute $\mu\alpha$

Division $\alpha \in \mathbf{R}, \beta \in \mathbf{R} \longrightarrow$ Compute $q = \lceil \alpha / \beta \rceil$

- Neglect the Gaussian Elimination Step.
- Use Nonnegative Integer Combination Instead of Convex Combination.
- Choose a Step Length Appropriately.

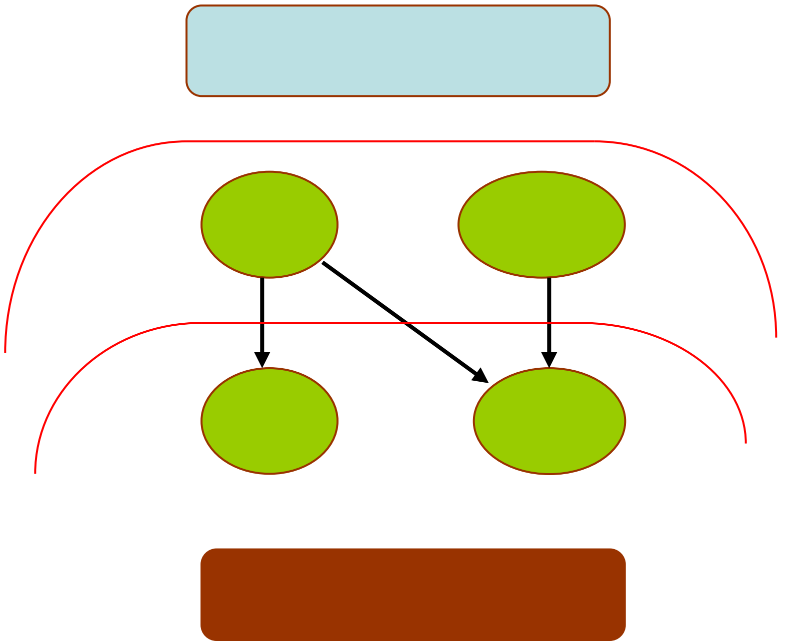
Finding All Minimizers

X, Y : Minimizer $\Rightarrow$ $X \cup Y, X \cap Y$: Minimizer

The set of minimizers forms a distributive lattice.

[G.Birkhoff]

Any distributive lattice can be represented as the set of ideals of a partial ordered set.



Partial Order of an Extreme Base

$y \in B(f)$: Extreme

Bixby, Cunningham, Topkis (1985)

$v_1 \quad v_2 \quad \cdot \cdot \cdot \quad v_n$

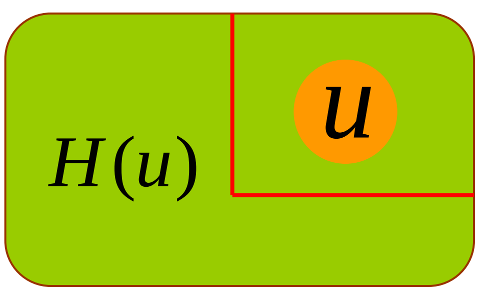
X : Tight $y(X) = f(X)$

$y(L(v)) = f(L(v)), \ \forall v \in V$

Represent all tight sets

$\longrightarrow G(y)$

v



$H(u)$: Maximal Ideal Excluding u.

Test if $H(u) \cup \{v\}$ is tight.

If not, then $v \longrightarrow u$.

Finding All Minimizers

Extreme Base

$$y_i \in B(f)$$

Partial Order (DAG)

$$G(y_i)$$

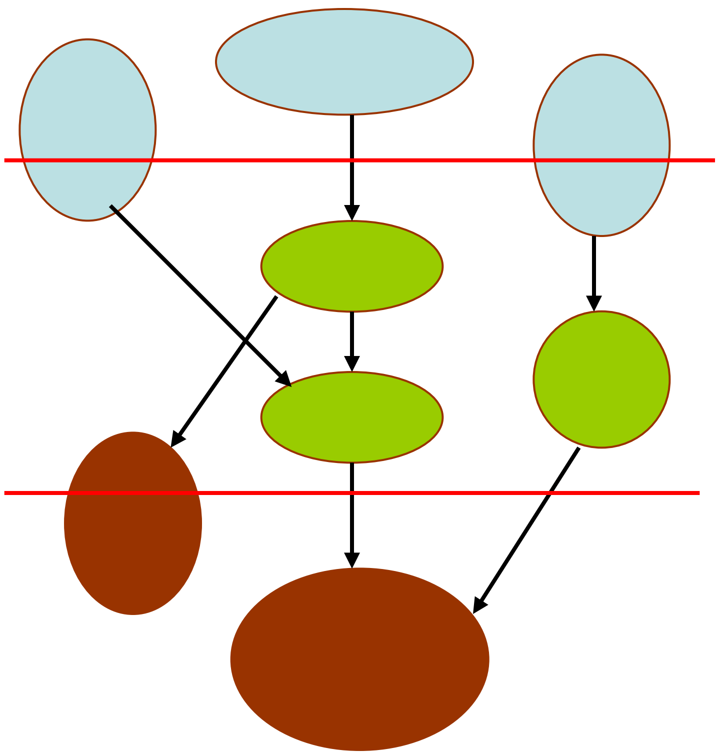
Convex Combination

$$x = \sum_{i \in I} \lambda_i y_i$$

$$(\lambda_i > 0,\ \forall i \in I)$$

$G(x)$: Superposition of $G(y_i)$

SCC Decomposition



$x(v) > 0$

$x(v) = 0$

$x(v) < 0$