

Extreme Compass and Dynamic Multi-Armed Bandits for Adaptive Operator Selection

Jorge Maturana¹, Álvaro Fialho², Frédéric Saubion¹,
Marc Schoenauer^{2,3} and Michèle Sebag^{2,3}

¹LERIA, Université d'Angers, Angers, France

²Microsoft Research – INRIA Joint Centre, Orsay, France

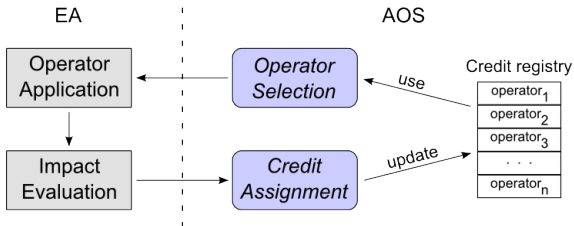
³Project-team TAO, LRI & INRIA Saclay, Orsay, France

IEEE Congress on Evolutionary Computation - CEC
May 19th, 2009

Parameter Control - Adaptive Operator Selection

Objective

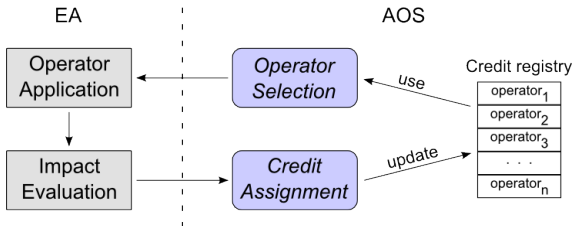
Autonomously select the operator to be applied amongst available ones, based on its impact in the past.



Parameter Control - Adaptive Operator Selection

Objective

Autonomously select the operator to be applied amongst available ones, based on its impact in the past.



- Operator Selection: Dynamic Multi-Armed Bandits [Fialho et al., 2008]
- Credit Assignment: Compass [Maturana and Saubion, 2008]

AOS: A (kind of) Multi-Armed Bandit problem

Multi-Armed Bandits



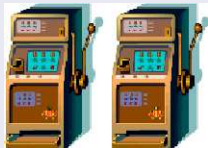
At time t , gambler plays arm j

$$\text{reward at } t : r_t = \begin{cases} 1 & \text{with prob} = p_j \\ 0 & \text{with prob} = 1 - p_j \end{cases}$$

Goal: maximize cumulated reward

AOS: A (kind of) Multi-Armed Bandit problem

Multi-Armed Bandits



At time t , gambler plays arm j

$$\text{reward at } t : r_t = \begin{cases} 1 & \text{with prob} = p_j \\ 0 & \text{with prob} = 1 - p_j \end{cases}$$

Goal: maximize cumulated reward

State-of-the-art: UCB1 [Auer et al., 2002]

- *Be optimistic in face of the unknown*
- At time t , choose arm j maximizing:

$$\hat{r}_{j,t} + \sqrt{\frac{2 \log \sum_k n_{k,t}}{n_{j,t}}}, \text{ where } \begin{cases} \hat{r}_{j,t} & \text{empirical reward for arm } j \\ n_{j,t} & \text{chosen times for arm } j \end{cases}$$

AOS with Multi-Armed Bandits: the true story

Scaling

MAB framework: $\hat{r}_{j,t} \in [0, 1]$;

AOS framework: $\hat{r}_{j,t} \in [a, b]$ (e.g. fitness improvement, diversity)

UCB1's EvE balance is broken, **Scaling** is needed:

$$\hat{q}_{i,t} = \hat{r}_{j,t} + c \sqrt{\frac{2 \log \sum_k n_{k,t}}{n_{j,t}}}$$

AOS with Multi-Armed Bandits: the true story

Scaling

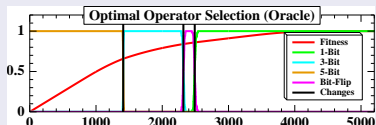
MAB framework: $\hat{r}_{j,t} \in [0, 1]$;

AOS framework: $\hat{r}_{j,t} \in [a, b]$ (e.g. fitness improvement, diversity)

UCB1's EvE balance is broken, **Scaling** is needed:

$$\hat{q}_{i,t} = \hat{r}_{j,t} + c \sqrt{\frac{2 \log \sum_k n_{k,t}}{n_{j,t}}}$$

AOS is a dynamic context



UCB1: too long to recover.

Page-Hinkley test [Page, 1954]

- Detect change (**threshold γ**);
- Restart the MAB.



AOS with Multi-Armed Bandits: the true story

Scaling

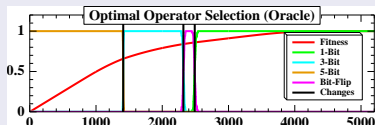
MAB framework: $\hat{r}_{j,t} \in [0, 1]$;

AOS framework: $\hat{r}_{j,t} \in [a, b]$ (e.g. fitness improvement, diversity)

UCB1's EvE balance is broken, **Scaling** is needed:

$$\hat{q}_{i,t} = \hat{r}_{j,t} + c \sqrt{\frac{2 \log \sum_k n_{k,t}}{n_{j,t}}}$$

AOS is a dynamic context



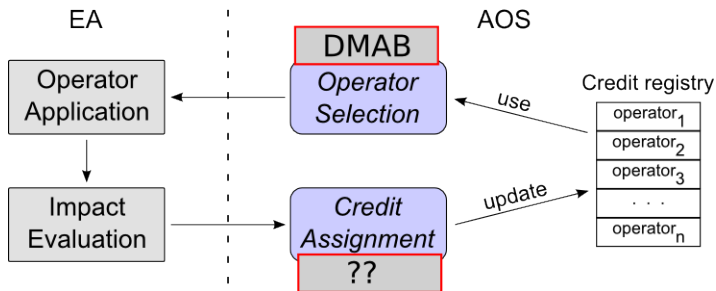
UCB1: too long to recover.

Page-Hinkley test [Page, 1954]

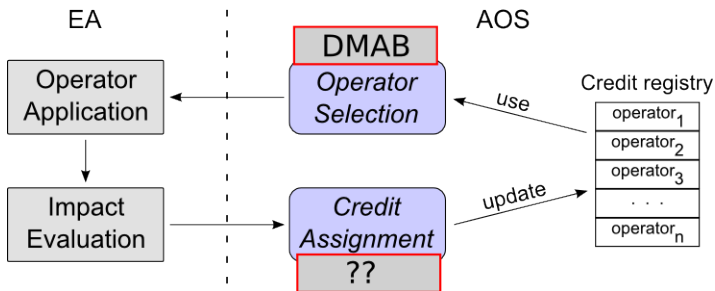
- Detect change (**threshold γ**);
- Restart the MAB.

DMAB: UCB1 + Scaling + Page-Hinkley

Credit Assignment

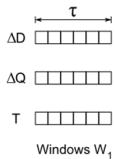


Credit Assignment

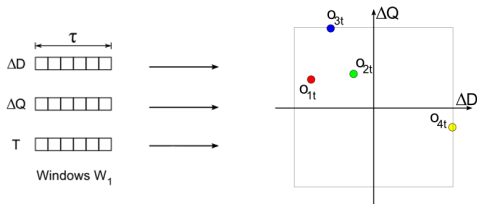


- Uni-modal problems
 - Fitness improvement
- Multi-modal problems (e.g. SAT)
 - Fitness improvement and Diversity are important
 - How to aggregate them?

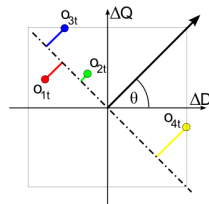
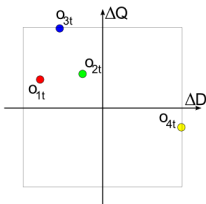
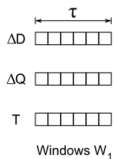
Credit Assignment: Compass



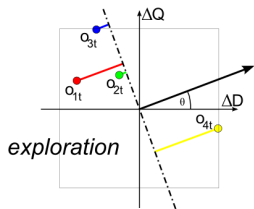
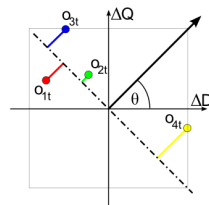
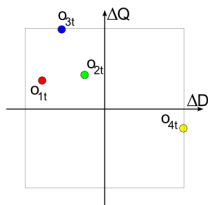
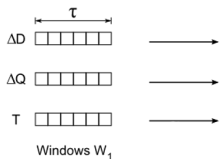
Credit Assignment: Compass



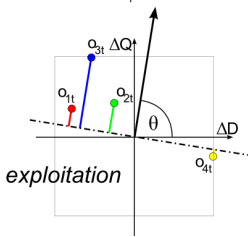
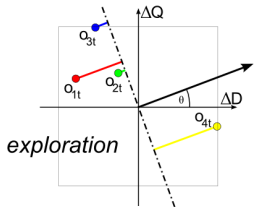
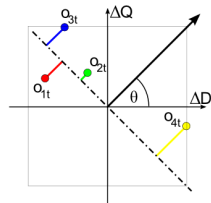
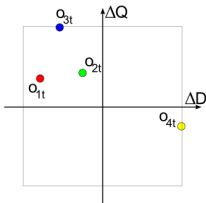
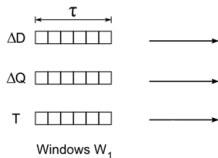
Credit Assignment: Compass



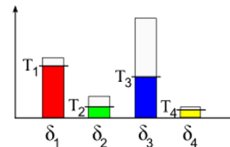
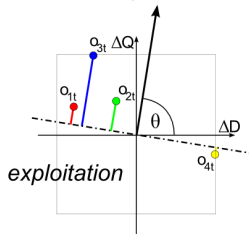
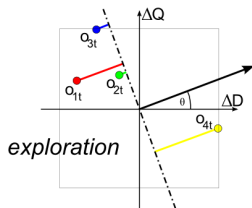
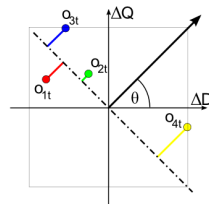
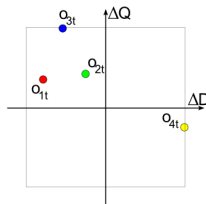
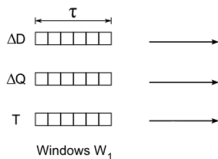
Credit Assignment: Compass



Credit Assignment: Compass



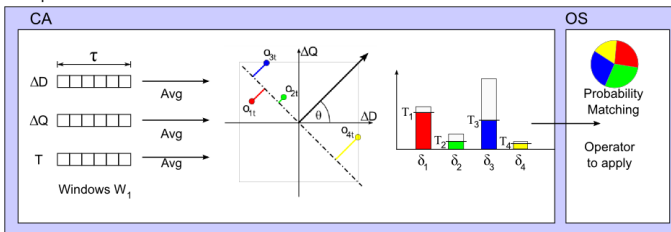
Credit Assignment: Compass



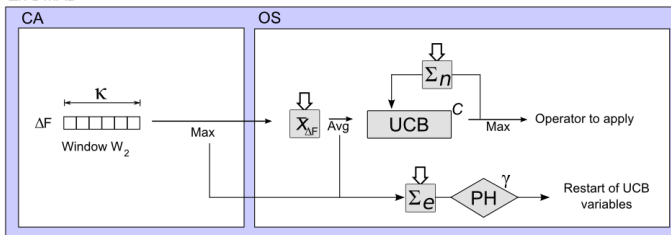
Reward divided by
Op. Application Time

Original AOS combinations by PPSN'08

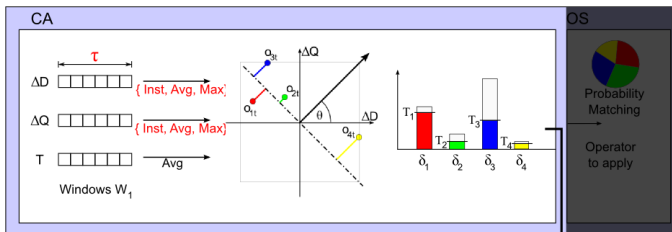
Compass



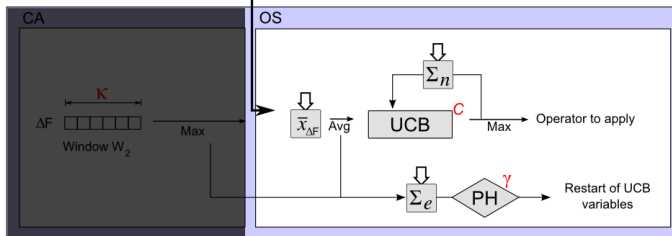
Ex-DMAB



Compass + Ex-DMAB = ExCoDyMAB



Compass + Ex-DMAB = ExCoDyMAB



Experimental Framework

ExCoDyMAB versus...

- Original Compass [Maturana and Saubion, 2008]
- Original DMAB [Fialho et al., 2008]
- Naive (uniform) operator selection

Experimental Framework

ExCoDyMAB versus...

- Original Compass [Maturana and Saubion, 2008]
- Original DMAB [Fialho et al., 2008]
- Naive (uniform) operator selection

Satisfiability (SAT) Problems

- Assign values to binary variables to satisfy a Boolean formula
- Objective: minimize $\#$ false clauses in the formula
- NP-complete, many real and artificial problems as instances
- Diverse set of fitness landscapes with different characteristics
- 22 instances used (from SATlib and SAT Race 2006)
 - Different families and levels of difficulty

Experimental Framework

Algorithm

- Representation: 1 bit per boolean variable
- Steady-state evolutionary algorithm
 - Population 3, Offspring replaces worst parent
 - Stopping at 5000 iterations

Experimental Framework

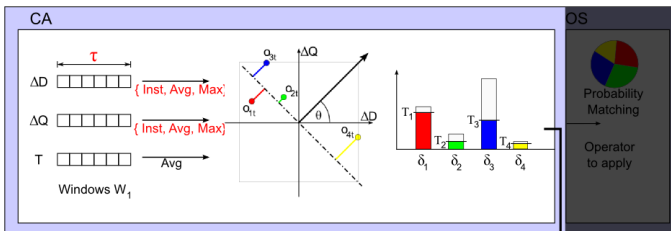
Algorithm

- Representation: 1 bit per boolean variable
- Steady-state evolutionary algorithm
 - Population 3, Offspring replaces worst parent
 - Stopping at 5000 iterations

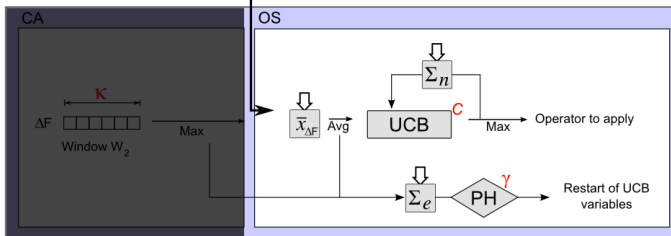
Operators

- Purpose: verify the ability of the AOS proposed
 - One-point Crossover
 - 1-Bit Hill Climbing
 - Contagion: false clauses of worst “contaminated” by the best
 - Tunneling: swaps variables while $\#$ true clauses grows
 - Bad Swap: swaps all variables that appears in false clauses
 - Wave: swaps variables in more false clauses and supporting less true ones

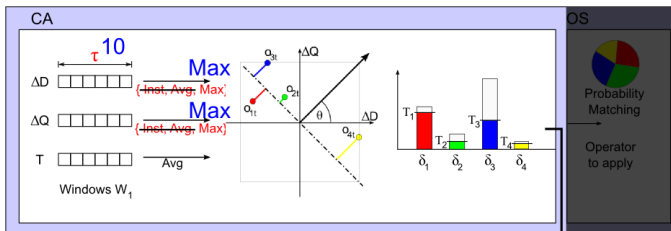
Compass + Ex-DMAB = ExCoDyMAB



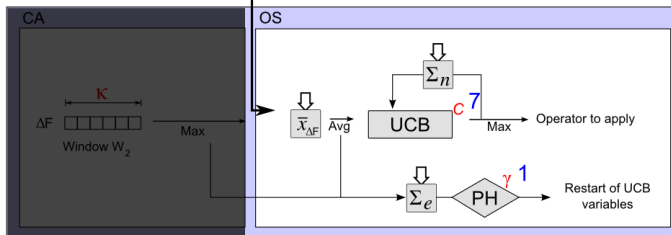
Compass + Ex-DMAB = ExCoDyMAB



Off-line Meta-Parameters Tuning: F-Race



Compass + Ex-DMAB = ExCoDyMAB



Comparative Results

- Best meta-parameter configuration for each technique
- 50 runs on each of 22 SAT instances
- “Wins - Losses” according to Student’s T-test 95%

	<i>Compass</i>	<i>DMAB</i>	<i>Naive</i>	$\sum dom$
<i>ExCoDyMAB</i>	18 - 2	21 - 0	22 - 0	59
<i>Compass</i>	–	9 - 9	22 - 0	8
<i>DMAB</i>	–	–	22 - 0	4

Conclusions

- AOS = Credit Assignment (CA) + Operator Selection (OS)
- Original Compass AOS: **sophisticated CA**, but **too simple OS**
- Original DMAB AOS: **just Δ fitness as CA**, but **engineered OS**
- ExCoDyMAB AOS: combines **good** elements of each one
- Results confirm the good synergetic effect of such combination
- Useful to any search technique with exploit/explore operators

Conclusions

- AOS = Credit Assignment (CA) + Operator Selection (OS)
 - Original Compass AOS: sophisticated CA, but too simple OS
 - Original DMAB AOS: just Δ fitness as CA, but engineered OS
 - ExCoDyMAB AOS: combines good elements of each one
 - Results confirm the good synergetic effect of such combination
 - Useful to any search technique with exploit/explore operators
-
- Preliminary (off-line) meta-parameter tuning still expensive
 - Deepen understanding on meta-parameters
 - Off-line? Self-Adaptation? Learning? Meta-Bandit?
 - Try with state-of-the-art SAT operators
 - Compare to known SAT solvers

References I



Auer, P., Cesa-Bianchi, N., and Fischer, P. (2002).

Finite-time analysis of the multiarmed bandit problem.

Machine Learning, 47(2/3):235–256.



Eiben, A. E., Michalewicz, Z., Schoenauer, M., and Smith, J. E. (2007).

Parameter Control in Evolutionary Algorithms.

In Lobo, F., Lima, C., and Michalewicz, Z., editors, *Parameter Setting in Evolutionary Algorithms*, chapter 2, pages 19–46. Springer Verlag.



Fialho, A., Da Costa, L., Schoenauer, M., and Sebag, M. (2008).

Extreme value based adaptive operator selection.

In G. Rudolph et al., editor, *Proc. PPSN'08*, pages 175–184. Springer.



Hartland, C., Baskiotis, N., Gelly, S., Teytaud, O., and Sebag, M. (2007).

Change point detection and meta-bandits for online learning in dynamic environments.

In *Proc. CAp'07*.

References II



Maturana, J. and Saubion, F. (2008).

A compass to guide genetic algorithms.

In G. Rudolph et al., editor, *Proc. PPSN'08*, pages 256–265. Springer.



Page, E. (1954).

Continuous inspection schemes.

Biometrika, 41:100–115.

Extreme Compass and Dynamic Multi-Armed Bandits for Adaptive Operator Selection

Jorge Maturana¹, Álvaro Fialho², Frédéric Saubion¹,
Marc Schoenauer^{2,3} and Michèle Sebag^{2,3}

¹LERIA, Université d'Angers, Angers, France

²Microsoft Research – INRIA Joint Centre, Orsay, France

³Project-team TAO, LRI & INRIA Saclay, Orsay, France

IEEE Congress on Evolutionary Computation - CEC
May 19th, 2009

SAT instances from SATlib

Problem	Sat?	# Vars.	# Clauses	Family
4blocks	Yes	758	47820	Blocks World Problem
aim	Yes	200	320	Random-3-SAT
f1000	Yes	1000	4250	Random-3-SAT
CBS	Yes	100	449	Controlled Backbone
Flat200	Yes	600	2237	Flat Graph Coloring
logistics	Yes	828	6718	Logistics Planning
medium	Yes	116	953	Randomly Generated
Par16	Yes	1015	3310	Parity Learning Problem
sw100-p0	Yes	500	3100	Morphed Graph Coloring
sw100-p1	Yes	500	3100	Morphed Graph Coloring
Uf250	Yes	250	1065	Phase Transition Region
Uuf250	No	250	1065	Phase Transition Region

SAT instances from SAT 2006 Race

Problem	Sat?	# Vars.	# Clauses	Family
Color*	No	1444	119491	Chessboard Coloring
G125*	Yes	2125	66272	Graph Coloring
Goldb-heqc*	No	5980	35229	Randomly Generated
Grieu-vmc	Yes	729	96849	Randomly Generated
Hoons-vbmc*	No	8503	25116	Randomly Generated
Schup	No	14809	48483	Randomly Generated
Simon*	No	2424	14812	Randomly Generated
Manol-pipe	Yes	14052	41596	Pipelined Machine Verification
Velev-eng*	No	6944	66654	Pipelined Machine Verification
Velev-sss*	No	1453	12531	Pipelined Machine Verification

Comparative Results I

Method Problem	<i>ExCoDyMAB</i> (C)	<i>Compass</i>	<i>Ex-DMAB</i>	<i>Uniform Choice</i>
4blocks	2.8 (0.9)	6 (0.9)	6.2 (0.9)	13.4 (0.6)
aim	1 (0)	1 (0)	1.2 (0.3)	3.6 (1.8)
f1000	10.3 (2.3)	30.9 (6.2)	16.4 (2.6)	55.8 (8.6)
CBS	0.6 (0.6)	0.4 (0.5)	1 (0.9)	7 (2.7)
Flat200	7.2 (1.7)	10.6 (2.1)	10.7 (2.2)	37.7 (5.5)
logistics	6.5 (1.3)	7.6 (0.5)	8.8 (1.5)	17.9 (4.1)
medium	1.5 (1.5)	0 (0)	1.8 (1.6)	8.8 (3.4)
Par16	15.2 (3.1)	64 (10.2)	24.1 (5.7)	131.1 (14.5)
sw100-p0	9.2 (1.2)	12.8 (1.4)	12.5 (1.7)	25.9 (3.4)
sw100-p1	0 (0)	0.5 (0.6)	1.1 (0.8)	11.3 (3.5)
Uf250	0.9 (0.7)	1.8 (0.9)	1.7 (0.8)	9.1 (3.3)
Uuf250	2.5 (1)	4.5 (1.2)	3.1 (1.1)	12.7 (3.2)

Comparative Results II

Method Problem	<i>ExCoDyMAB</i> (C)	<i>Compass</i>	<i>Ex-DMAB</i>	<i>Uniform Choice</i>
Color	48 (2.5)	61.3 (2.2)	49.3 (3.4)	80.4 (6.6)
G125	8.8 (1.3)	20.6 (2)	13.5 (1.7)	28.8 (4.6)
Goldb-heqc	72.9 (8.5)	112.2 (15.2)	133.2 (15.9)	609.7 (96.2)
Griev-vmc	16.7 (1.7)	15.2 (1.7)	19.6 (1.8)	24.1 (3.3)
Hoons-vbmc	69.7 (14.5)	268.1 (44.6)	248.3 (24.1)	784.5 (91.9)
Manol-pipe	163 (18.9)	389.6 (37.2)	321 (38.1)	1482.4 (181.5)
Schup	306.6 (26.9)	807.9 (81.8)	623.7 (48.5)	1639.5 (169.9)
Simon	29.6 (3.3)	43.5 (2.7)	35.3 (6.3)	72.6 (11.3)
Velev-eng	18.3 (5.2)	29.5 (7.3)	118 (37.1)	394 (75.8)
Velev-sss	2 (0.6)	4.6 (1)	5.9 (3.9)	62.7 (25.2)