Flexible FOND Planning under Explicit Fairness Assumptions

Ivan D. Rodriguez,[1] Blai Bonet,[1] Sebastian Sardiña,[2] Hector Geffner[1,3]

[1]Universitat Pompeu Fabra, Barcelona, Spain [2]RMIT University, Melbourne, Australia [3]Institució Catalana de Recerca i Estudis Avançats (ICREA)

arXiv:2103.08391

Motivation and Contribution

- **Goal**: combine & extend in a uniform setting:
  1. Fair FOND Planning: strong cyclic; proper policies of MDPs.
  2. Adversarial FOND Pl.: strong; bounded # of steps.
  3. Qualitative Numerical Pl.: used in generalized planning.
- **Problem**: their models and solvers have *implicit* fairness assumptions, thus they cannot be combined.
- **Solution**:
  1. define FOND$^+$ planning, making these assumptions explicit;
  2. define SIEVE$^+$ procedure to test FOND$^+$ solutions; and
  3. implement FOND-ASP planner for FOND$^+$ in CLINGO.

Long paper (proofs + ext. discussions) in **arXiv:2103.08391**

FOND Model and Problem

A **FOND model** is a tuple $M = \langle S, s_0, S_G, Act, A, F\rangle$, where:

- S is a finite set of states;
- $s_0 \in S$ is the initial state;
- $S_G \subseteq S$ is a non-empty set of goal states;
- Act is a set of actions;
- $A(s) \subseteq Act$ is the set of actions applicable in state s; and
- $F(a, s)$ is the set of successor states when action a is executed in state s.

A **FOND problem** $P = \langle At, I, Act, G\rangle$ is a compact description of a FOND model $M(P)$.

Action effects in FOND planning can be deterministic of the form E_i, or **non-deterministic** of the form $oneof(E_1, \ldots, E_n)$.

FOND Solutions

Policy π: partial function mapping *non-goal* states into actions.

π-trajectory: possibly infinite state trajectory $s_0, s_1, s_2, \ldots$ compatible with policy π for P, where $s_{i+1} \in F(\pi(s_i), s_i)$, for $i \geq 0$.

Maximal π-trajectory: is infinite, or ends in a state where $\pi(s)$ is undefined or not applicable.

Solutions:

1. π is a **strong cyclic solution** of P if every reachable state is connected to a goal state with the policy.
2. π is a **strong solution** of P if all maximal **π-trajectory** are finite and end in a goal state.

Fairness in FOND planning

*A **policy** π **solves problem** P when all maximal **fair** trajectories compatible with π reach the goal*, provided fairness is defined follows:

- Strong-cyclic planning: all finite trajectories are fair; and infinite trajectories are fair *iff* all recurrent states s are directly followed by each successor $s' \in F(\pi(s), s)$ an infinite number of times.
- Strong planning: all trajectories are deemed fair.

Qualitative Numeric Planning Planning

A **QNP** $Q = \langle At, V, I, O, G\rangle$ extends a STRIPS problem with a set V of *numerical variables* X that can appear in:

- effects as *qualitative* increments $X\uparrow$ and decrements $X\downarrow$; and
- literals $X = 0$ or $X > 0$.

Unlike in numerical planning, plan existence for QNPs is decidable!

A QNP Q **defines a standard FOND** problem $P = T_D(Q)$ where:

- Each $n \in V$ is replaced by a boolean p_n that stands for $n = 0$.
- Literals $n = 0$ and $n > 0$ are replaced by p_n and $\neg p_n$, respectively.
- Effects $n\uparrow$ are replaced with effects $\neg p_n$.
- Effects $n\downarrow$ are replaced with effects $oneof(p_n, \neg p_n)$.

QNP Example: Clearing a Block

Clearing a block x in a Blocksworld instance.

- Boolean variable H: true if holding a block.
- Numerical variable n: number of blocks above x.
- Initially $I = \{\neg H, n > 0\}$. Goal $G = \{n = 0\}$.

Actions:

- $Pick\text{-}above\text{-}x = \langle \neg H, n > 0; H, n\downarrow\rangle$
- $Put\text{-}above\text{-}x = \langle H; \neg H, n\uparrow\rangle$
- $Pick\text{-}other = \langle \neg H; H\rangle$
- $Putaway = \langle H; \neg H\rangle$

Solution: if $\neg H$ then $Pick\text{-}above\text{-}x$. if H then $Putaway$

QNP termination: Sieve

Srivasta et al. 2011:

Sieve

Let π be a policy for the FOND problem $P = T_D(Q)$ associated with QNP Q. The policy π **terminates** in P iff every state s that is reachable by π in P terminates, where a state s **terminates** iff:

1. there is no cycle on node s (i.e., no path from s to itself);
2. every cycle on s contains a state s' that **terminates**; or
3. $\pi(s)$ decrements a variable x, and every cycle on s containing a state s' for which $\pi(s')$ increments x, also contains a state s'' that **terminates**.

Theorem

A policy π is a solution to a QNP Q **iff** π is a strong-cyclic solution of $P = T_D(Q)$ that terminates.

FOND$^+$ Problems

Definition

A FOND$^+$ problem $P_c = \langle P, C\rangle$ is a FOND problem P extended with a set C of **(conditional) fairness assumptions** of the form A_i/B_i, with $i = 0, \ldots, n$, and where each A_i is a set of **non-deterministic actions** in P and each B_i is a set of actions in P disjoint from A_i.

FOND+ Fairness. Meaning of $A/B \in C$: If a state trajectory contains infinite occurrences of actions $a \in A$ in a state s, and *finite* occurrences of actions from B, then s must be immediately followed by each $s' \in F(\pi(s), s)$ an infinite number of times.

Solutions

A policy π **solves** the FOND$^+$ problem $P_c = \langle P, C\rangle$ if all the maximal **π-trajectory** that are **fair** reach the goal.

FOND$^+$ Fairness

Theorem

The **strong-cyclic solutions** of a FOND problem P are the solutions of the FOND$^+$ problem $P_c = \langle P, \{A/\emptyset\}\rangle$, where A is the set of all the non-deterministic actions in P.

Theorem

The **strong solutions** of a FOND problem P are the solutions of the FOND$^+$ problem $P_c = \langle P, \emptyset\rangle$.

Theorem

The solutions of a **QNP problem** Q are the solutions of the FOND$^+$ problem $P_c = \langle P, C\rangle$ where $P = T_D(Q)$ and C is the set of fairness assumptions A_i/B_i, one for each numerical variable x_i in Q, such that:

1. A_i contains all the actions in P that decrement x_i
2. B_i contains all the actions in P that increment x_i.

FOND$^+$ termination: Sieve$^+$

Sieve$^+$

Let π be a policy for the FOND$^+$ problem $P_c = \langle P, C\rangle$.
State s in P **terminates** iff

1. s is a goal state;
2. s is **fair** and some state $s' \in F(\pi(s), s)$ **terminates**; or
3. s is **not fair**, all states $s' \in F(\pi(s), s)$ **terminate**, and $F(\pi(s), s)$ is non-empty.

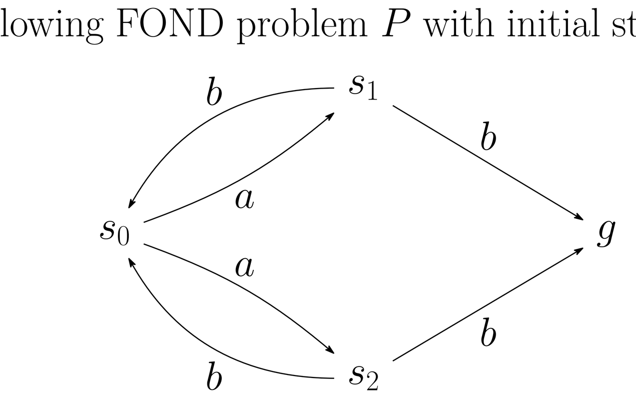
Here, s is **fair** if for some A_i/B_i in C, $\pi(s) \in A_i$, and every path that connects s to itself and that contains a state s' with $\pi(s') \in B_i$, also contains a state s'' that **terminates**.

Theorem

A policy π solves the FOND$^+$ problem $P_c = \langle P, C\rangle$ iff all the states s that are reachable by π terminate.

FOND$^+$ Example

Given the following FOND problem P with initial state s_0 and goal g:



FOND$^+$ problems $P_{c_i} = \langle P, C_i\rangle$ are solvable (✓) or unsolvable (✗):

- ✗ $C_1 = \{\}$; a and b are adversarial.
- ✓ $C_2 = \{a, b\}$; a and b are fair.
- ✗ $C_3 = \{a\}$; a is fair and b is adversarial.
- ✓ $C_4 = \{b\}$; b is fair and a is adversarial.
- ✗ $C_5 = \{a/b\}$; fairness for a is false; b adversarial.
- ✗ $C_5 = \{a/b\}$; a is conditionally fair on b; b adversarial.
- ✗ $C_6 = \{a, b/a\}$; QNP like: $a : x_1\downarrow, x_2\uparrow$ and $b : x_2\downarrow$.
- ✓ $C_7 = \{b, a/b\}$; QNP like: $b : x_1\downarrow, x_2\uparrow$ and $a : x_2\downarrow$.
- ✗ $C_8 = \{a/b, b/a\}$; QNP like: $a : x_1\downarrow, x_2\uparrow$ and $b : x_2\downarrow, x_1\uparrow$.

FOND-ASP: An ASP-based FOND$^+$ Planner

```
% policy, edges, and connectedness
{ pi(S,A) : ACTION(A) } = 1 :- STATE(S), not GOAL(S).
edge(S,T) :- pi(S,A), TRANSITION(S,A,T).

connected(S,T) :- edge(S,T).
connected(S,T) :- connected(S,X), edge(X,T), S != X.

blocked(S,T) :- STATE(S), STATE(T), not connected(S,T).
blocked(S,T) :- connected(S,T), terminate(S).
blocked(S,T) :- connected(S,T), terminate(T).
blocked(S,T) :- connected(S,T),
                blocked(X,T) : edge(S,X), connected(X,T).

fair(S) :- pi(S,A), ASET(I,A), blocked(X,S) : pi(X,B),
  BSET(I,B), not blocked(S,X).

% terminating states
terminate(S) :- GOAL(S).
terminate(S) :- fair(S), edge(S,T), terminate(T).
terminate(S) :- not fair(S), edge(S,_), terminate(T) : edge(S,T).

% reachable states must terminate
:- reachable(S), not terminate(S).
reachable(S) :- INITIAL(S).
reachable(S) :- reachable(X), not GOAL(X), edge(X,S).
```

Experiments

QNP problem	#states	QNP2FOND FOND-SAT	QNP2FOND PRP	STRIX	FOND-ASP
qnp2-02	8	0.20	0.18	2.33	0.00
qnp2-03	16	1.77	0.30	2.31	0.01
qnp2-04	32	10.00	0.58	14.25	0.04
qnp2-05	64	50.24	1.15	885.37	0.20
qnp2-06	128	302.80	2.53	—	1.26
qnp2-07	256	1,969.35	4.02	—	7.14
qnp2-08	512	—	6.96	—	54.37
qnp2-09	1,024	—	13.22	—	***
qnp2-10	2,048	—	21.94	—	***

FOND$^+$ problem	$f01$ (unsolvable) #states	$f01$ STRIX	$f01$ FOND-ASP	$f11$ (solvable) #states	$f11$ STRIX	$f11$ FOND-ASP
qnp2-$f$xx-02	8	3.22	0.00	32	5.85	0.04
qnp2-$f$xx-03	16	2.25	0.01	64	8.16	0.21
qnp2-$f$xx-04	32	11.38	0.04	128	236.89	1.55
qnp2-$f$xx-05	64	873.09	0.21	256	—	15.45
qnp2-$f$xx-06	128	—	1.25	512	—	46.67
qnp2-$f$xx-07	256	—	12.13	1,024	—	***
qnp2-$f$xx-08	512	—	39.56	2,048	—	***
qnp2-$f$xx-09	1,024	—	***	4,096	—	***
qnp2-$f$xx-10	2,048	—	***	8,192	—	***

Wrap Up, Conclusions, and Future Work

- Unified treatment of strong, strong cyclic, QNP planning, and beyond, through explicit fairness assumptions A/B where A and B are sets of actions
- FOND$^+$ planning more expressive than existing FOND planning models, but less expressive and less complex than LTL planning
- Simple but effective **flat** FOND+ planner in ASP/CLINGO
- Future work: factored FOND$^+$ planner, scalable and sound, but not complete