Optimal Strategies in Request-Response Games

Nico Wallmeier

Lehrstuhl Informatik 7
RWTH Aachen

April 11, 2007

Outline

Request-Response Winning Condition

- Game graph $G = \{Q, E\}$
- *Request-Response winning condition*:
  - Sets $P_i, R_i \subseteq Q$ for $1 \leq i \leq r$
  - (P_i, R_i) is called *RR pair* or *RR condition*
  - Winning condition given by
    - $\bigwedge_{i=1}^{r}$ if play π reaches P_i at time j, then also R_i at a time $j' \geq j$
    - "whenever a state in P_i is visited, then eventually a state in R_i is visited"
    - Or in LTL: $\bigwedge_{i=1}^{r} \mathsf{G}(P_i \rightarrow \mathsf{F}\ R_i)$
  - Similar to Streett games

Reduction to Büchi Games

Theorem

RR games are reducible to Büchi games, involving a blow-up from n to $nr2^{r+1}$ states if r RR conditions are involved.

Idea

Keep in mind (encoding in an expanded game graph):

- Which requests are not yet fulfilled
- Which one should be fulfilled next (progressing cyclically, thereby visiting final states)

Motivation

- RR winning condition only ensures that R_i is visited eventually (there is no specification when)
- In the context of controller synthesis: we want/need *optimal strategies*
- Example: lift controller, call center, ...

Plan

- Define value of plays
- Define optimal strategies in RR games
- Crucial point: There are finite-state optimal winning strategies
- Key for computing optimal strategies, use results on mean-payoff games

RR-Conditions

Definition

A condition (P_i, R_i) in play ρ of a RR-game at time t is called

(a) active, iff $\exists t' \leq t : \rho(t') \in P_i \wedge \forall s \in [t', t] : \rho(s) \notin R_i$

$$active(\rho, t, i) = \begin{cases} 1 & \text{in play } \rho \text{ condition } i \text{ at time } t \text{ is active} \\ 0 & \text{otherwise} \end{cases}$$

(b) activated, iff $\neg active(\rho, t-1, i) \wedge active(\rho, t, i)$

$$newActive(\rho, t, i) = \begin{cases} 1 & active(\rho, t, i) \wedge \neg active(\rho, t-1, i) \\ 0 & \text{otherwise} \end{cases}$$

(c) fulfilled, iff $active(\rho, t-1, i) \wedge \rho(t) \in R_i$

Waiting time

Definition (Waiting time)

The waiting time of a condition (P_i, R_i) in a play ρ of RR-game at time t is:

$$WZ(\rho, t, i) = \begin{cases} 0 & \neg active(\rho, t, i) \\ t - t' + 1 & \exists t' \leq t : newActive(\rho, t', i) \wedge \\ & \forall s \in [t', t] : active(\rho, s, i) \end{cases}$$

Definition (Waiting bound)

The waiting bound $wb(\rho, i)$ of i-th condition (P_i, R_i) in a play ρ is the maximal waiting time if it exists, otherwise ∞.

Waiting time (2)

Note

The waiting bound $wb(\rho, i)$ is infinite, if

- the condition is not fulfiled at some time in ρ, or
- the sequence of the waiting times is unbounded.

Note

In a play ρ a condition i is called unbounded, if

- the waiting time for this pair i is unbounded, and
- after every activation a fulfilment follows in ρ

Waiting time vector

Definition (Waiting time vector)

A play ρ of a RR-game with r pairs can be expanded with the waiting time vector (wtv) $\begin{pmatrix} WZ(\rho, t, 1) \\ \vdots \\ WZ(\rho, t, r) \end{pmatrix} \in \mathbb{N}^r$ at time t.

Note

- A game graph of a RR-game can be expanded with the waiting time vectors - called G_r^+.
- The resulting game graph is not finite unless the waiting time for all pairs and all plays can be bounded.

Some Observations

Lemma (Bounded Waiting Time)

In a RR game with n states and r pairs a finite-state winning strategy ρ guarantees bounded waiting time for every play and every RR pair.

This Lemma does *not* hold in general, if

- the winning strategy is not finite-state
- the game graph is not finite

First Approach

Intuition: Average open requests

Definition (Value of a play)

Given a RR game with r pairs and a play ρ.

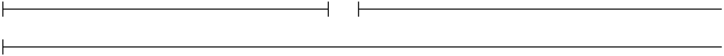
1. $a(\rho) = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{i=1}^{r} active(\rho, t, i)$

Discussion of $a(\rho)$

- Can not differentiate the following cases:
  - One or no fulfillment

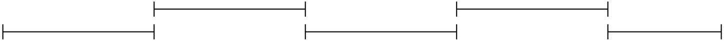
ρ_1:

ρ_2:



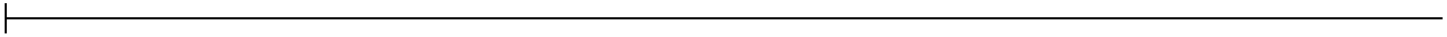
  - Alternation or no fulfillment

ρ_3:



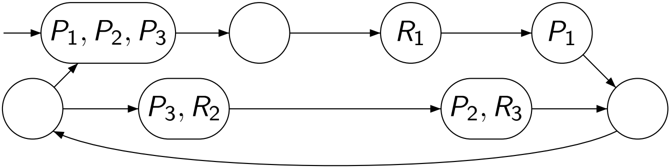
vs.

ρ_4:



Discussion of $a(\rho)$ - 2

- For achieving optimality, finite-state strategies do not suffice. One-player game:



Second Approach

Intuition: Average time requests are active

Definition (Value of a play - continued)

Given a RR game with r pairs and a play ρ.

2. $$t(\rho) = \lim_{n\to\infty} \frac{\sum_{t=1}^{n}\sum_{i=1}^{r} active(\rho,t,i)}{\sum_{t=1}^{n}\sum_{i=1}^{r} newActive(\rho,t,i)}$$

Lemma (Existence of the Limes)

If the play is played according to a finite-state winning strategy of player 0 and won by player 0, the Limes exists.

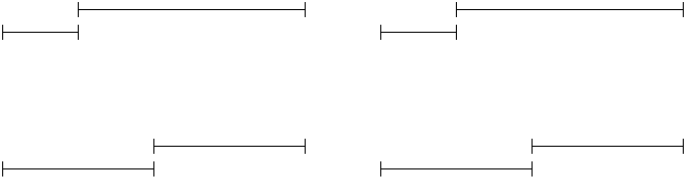
Discussion of $t(\rho)$

- Does not differentiate the following situation

ρ_5:

vs.

ρ_6:



Third Approach

Intuition: Punishment for longer waiting

Definition (Value of a play - continued)

3. $w(\rho) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{i=1}^{r} WZ(\rho, t, i)$

Lemma (Existence of the Limes)

If the play is played according to a finite-state winning strategy of player 0 and won by player 0, the Limes exists.

Definition (Value of a play segment)

$$w(\rho, t_1, t_2) = \frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} \sum_{i=1}^{r} WZ(\rho, t, i)$$

Discussion of $w(\rho)$

- Punishment for longer waiting
- The time to fulfill a condition is counted quadratically $(\sum_{i=0}^{n} i = \frac{n \cdot (n+1)}{2})$
- Therefore longer response times have a deeper impact
- Can distinguish the previous cases

Value of a strategy

$\rho(\sigma, \tau, q_0)$ is a unique play.

Definition (Value of a strategy)

The value of a strategy σ of player 0 from a state q_0 for a given RR game is defined as

$$w(\sigma, q_0) = \sup_{\tau \in \text{strategy of player 1}} w(\rho(\sigma, \tau, q_0))$$

Corollary (Existence of the Supremum)

For a given RR game, a finite-state winning strategy σ of player 0, and a initial state $q_0 \in W_0$, the supremum $w(\sigma, q_o)$ exists.

Optimality of a winning strategy

Definition

A winning strategy σ is *optimal* for a given RR game with initial state q_0 iff

$$w(\sigma, q_0) \leq w(\tau, q_0) \text{ for all strategies } \tau$$

Note

The optimality condition does not ensure a unique optimal strategy for a given RR game with initial state q_0.

Upper bound of a strategy value

Corollary

In a RR-game with n states and r pairs the strategy value of an optimal strategy σ is always:

$$w(\sigma, q_0) \leq \frac{n \cdot r^2 + r}{2} \text{ for } q_0 \in W_0$$

Proof.

- Büchi Reduction: a condition is atmost $n \cdot r$ steps active
- Value for a condition, which is $n \cdot r$ steps active: $\sum_{i=0}^{n \cdot r} i = \frac{n \cdot r \cdot (n \cdot r + 1)}{2}$
- Thus r conditions: $\frac{n \cdot r^2 \cdot (n \cdot r + 1)}{2}$
- The strategy value is smaller or equal: $\frac{1}{n \cdot r} \cdot \frac{n \cdot r^2 \cdot (n \cdot r + 1)}{2} = \frac{n \cdot r^2 + r}{2}$

Bounded waiting time

Theorem

There exists an optimal strategy in every RR game, which has a bounded waiting time for all RR-pairs.

Assumption for special case: Only one RR-pair is unbounded and the strategy σ is optimal.

Motivation: Decrease large value of first component of wtv by changing the play.

Consider a play ρ on G_r^+, which

- is played according to the strategy σ,
- won by player 0,
- the first RR-pair (P_1, R_1) is unbounded and
- the i-th pair is bounded by s_i for $2 \leq i \leq r$

.

Bounded waiting time (2)

Create a new play ρ' by checking for each position t, if a position t' exists, such that

1. the first pair is constantly active between t and t'
2. $\rho(t) = \rho(t') \in Q_0$
3. $\rho(t+1) \neq \rho(t'+1)$ (different successor states)
4. the waiting time vectors are equal at both positions except the first component
5. $w(\rho, t, t') > w(\rho)$

For waiting times $w > n \cdot \Pi_{i=2}^{r} s_i$ of the first pair, the first fourth conditions are fulfilled.

If such a maximal t' exists, the new play ρ' arises thereby that at position t it will be played that way as on position t' in ρ, otherwise the play will not be changed.

Bounded waiting time (3)

Easy to see:

- all pairs in the play ρ' are bounded
- $w(\rho') < w(\rho)$

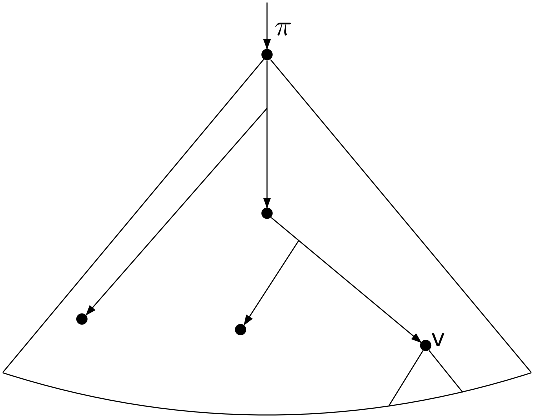
Definition

The game tree t_π^σ of a play prefix $\pi = \pi(0) \dots \pi(n)$ and a strategy σ of player 0 has at its root the node $\pi(n)$. All possible plays according to σ from prefix π are the paths in the tree.

In order to define a new winning strategy σ' for a play prefix π, take a look at the strategy tree t_π^σ, where all paths are cut after the first visit of R_1.
Note: t_π^σ is finite, if $\pi(0) \in W_0$

Bounded waiting time (4)

Game tree t_{π}^{σ}:



Bounded waiting time (5)

- Consider all nodes in the tree, which are identical to the root execpt of the first component of the wtv.
- Choose node v out of this set with maximal first component of the wtv (if not unique, choose one randomly)

New winning strategy σ' is defined as follows:

- if $w(\tau, t, v) > w(\sigma, \tau(0))$: strategy σ' chooses for the play prefix τ the successor of v, first component of wtv is set to the value of v
- otherwise the strategy σ' behaves like σ

Bounded waiting time (6)

Conditions for the general case:

1. an unbounded pair is constantly active during t and t'
2. $\rho(t) = \rho(t') \in Q_0$
3. $\rho(t+1) \neq \rho(t'+1)$ (different successor states)
4. all components of the waiting time vector at position t' are equal or bigger than on position t
5. $w(\rho, t, t') > w(\rho)$

It is then possible to derive a waiting bound for all conditions, which is called $w_{max}(n, r)$.

Memory of optimal strategies

Corollary

Finite memory is sufficient for an optimal strategy in a RR-game.

Proof.

- Optimal strategies are only depending on the waiting time of the RR-pairs.
- As show before these waiting times are bounded by an optimal strategy.
- Therefore finite memory is sufficient for optimal strategies in RR games.

□

Plan

Theorem

Optimal strategies for both players in a RR game are computable.

Idea

- Expand the game graph by the waiting time vectors which are bounded by $w_{max}(n, r) = w_{max}$ for all pairs.
- Reduce the RR game to a mean-payoff game

Mean-Payoff Games

Definition (Mean-Payoff Game, EM79)

Let $G = (Q, E)$ be a game graph and $wf : E \to \{-W, \dots, 0, \dots, W\}$ be a function that assigns an integral weight to each edge. Player 0 wants to maximize $\liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} wf(\rho_i)$, whereas player 1 wants to minimize $\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} wf(\rho_i)$.

Theorem (ZP96)

In a mean-payoff game $G = (Q, E)$ with a weight function $wf : E \to \{-W, \dots, W\}$ positional optimal strategies for both players are computable in time $\mathcal{O}(|Q|^4 \cdot |E| \cdot \log(|E|/|Q|) \cdot W)$.

Expanded game graph

- $Q_i' := Q_{1-i} \times \{0, \ldots, w_{max}+1\}^r$
- $((q, i_1, \ldots, i_r), (q', i_1', \ldots, i_r')) \in E' :\Leftrightarrow$
  - $(q, q') \in E$
  - $i_k' = \begin{cases} i_k + 1, & \text{if } 1 \leq i_k \leq w_{max} \wedge q' \notin R_k \\ 0, & \text{if } q' \in R_k \\ 1, & \text{if } i_k = 0 \wedge q' \in P_k \\ i_k, & \text{otherwise} \end{cases}$
- $f : Q \rightarrow Q'$ as follows:

$$f(q) = (q, i_1, \ldots, i_r) \text{ with } i_k = \begin{cases} 1 & \text{if } q \in P_k \\ 0 & \text{otherwise} \end{cases}$$

Reduction to Mean-Payoff Games

- weight function $wf : E' \longrightarrow \{0, \ldots, r^2 \cdot w_{max} + r\}$ as follows:

$$wf((p, \begin{pmatrix} w_1 \\ \vdots \\ w_r \end{pmatrix}), (q, \begin{pmatrix} w'_1 \\ \vdots \\ w'_r \end{pmatrix})) =$$

$$\sum_{k=1}^{r} (open(w'_k) \cdot w'_k + failed(w'_k) \cdot (r \cdot w_{max} + 1))$$ with

$$open(i) = \begin{cases} 1 & \text{if } 0 \leq i \leq w_{max} \\ 0 & \text{otherwise} \end{cases}$$ and

$$failed(i) = \begin{cases} 1 & \text{if } i = w_{max} + 1 \\ 0 & \text{otherwise} \end{cases}$$

- Positional optimal strategies in mean-payoff game according to Zwick/Paterson provide us the optimal finite-state winning strategies in the given RR game.

Conclusion

- We discussed different approaches of a play value for RR games
- Definition of optimal strategies in RR games
- Computation of optimal strategies