موسسه بابان

انتشارات بابان و انتشارات راهیان ارشد

درس و کنکور ارشد

سیستم عامل

(مدیریت فرآیندها و نخ‌های هم‌روند)

ویژه‌ی داوطلبان کنکور کارشناسی ارشد مهندسی کامپیوتر و IT

براساس کتب مرجع

آبراهام سیلبرشاتز، ویلیام استالینگز و اندرو اس تنن‌بام

ارسطو خلیلی‌فر

۷ فرآیندهای همروند

مقدمه

در اغلب سیستم‌های امروزی، تعدادی از فرآیندها یا نخ‌ها به صورت همروند بر روی یک پردازنده و یا به صورت موازی بر روی چندین پردازنده اجرا می‌شوند. در سیستم‌های چند برنامگی و چند پردازنده‌ای، همروندی فرآیندها و نخ‌ها، یک پدیده‌ی عادی به شمار می‌آید.

فرآیندهای هم روند و همکار، به ارتباط با یکدیگر نیاز دارند. آن‌ها برای دستیابی به یک هدف مشترک، نیازمند همکاری، هماهنگی، تبادل داده و استفاده از داده‌ها و سایر منابع مشترک هستند. بنابراین مدیریت اجرای همروند چند فرآیند بر روی یک پردازنده و اجرای موازی چند فرآیند بر روی چندین پردازنده، حائز اهمیت فراوان می‌باشد. این مدیریت باید به گونه‌ای باشد که اجرای یک فرآیند آسیبی به اجرای فرآیند(های) همکار دیگر نرساند.

در هنگام طراحی سیستم عامل، در زمینه‌ی ارتباط بین فرآیندها با سه مسأله‌ی اساسی زیر مواجه هستیم:

۱ ـ تبادل داده

گاهی یک فرآیند، به نتیجه‌ی محاسبات یک فرآیند دیگر نیاز دارد، بنابراین به یک مکانیسم برای ارتباط بین فرآیندها نیاز است. انواع مکانیزم‌های تبادل داده بین فرآیندها به روش‌های زیر است:

- حافظه مشترک

- فایل مشترک
- تبادل پیام (در انتهای فصل تشریح خواهد شد)
- لوله (نوعی شبه فایل در سیستم عامل یونیکس)

تبادل داده برای نخ‌ها ساده می‌باشد، زیرا نخ‌ها یک فضای آدرس مشترک دارند. نخ‌های متعلق به فرآیندهای جداگانه که در فضای آدرس متفاوت قرار دارند، در صورت نیاز به ارتباط باید از مکانیسم‌های ارتباط فرآیندها، استفاده کنند که نیازمند فراخوان‌های سیستمی وقت‌گیر هستند.

ناحیه‌ی بحرانی

اگر چند فرآیند قصد دسترسی به یک منبع مشترک را داشته باشند، قطعه کدی از هر فرآیند را که در آن به دستکاری این منبع مشترک می‌پردازد، ناحیه‌ی بحرانی می‌گویند.

نکته : در همه‌ی نواحی بحرانی، دسترسی به منبع مشترک، وجود دارد، اما عکس آن همیشه صادق نیست و هرگونه دسترسی به منبع مشترک باعث رقابت و ایجاد ناحیه‌ی بحرانی نمی‌شود.

منبع بحرانی

منبعی که توسط ناحیه‌ی بحرانی مورد دستیابی قرار می‌گیرد، منبع بحرانی نام دارد، مانند متغیرهای مشترک رقابت‌زا.

۲ ـ شرایط رقابتی (مسابقه)

هرگاه دو یا چند فرآیند همزمان با هم وارد ناحیه‌ی بحرانی (منبع مشترک) شوند، شرایط رقابتی پیش می‌آید. در شرایط رقابتی، نتیجه‌ی نهایی بستگی به ترتیب دسترسی‌ها دارد. در واقع فرآیندهای همکار بر هم اثر دارند و اینکه پردازنده، به چه ترتیبی و در چه زمان‌هایی بین آنها تعویض متن انجام دهد در ایجاد پاسخ نهایی اثرگذار خواهد بود. بنابراین علت شرایط رقابت **تعویض متن پردازنده** بین فرآیندهای همکار است.

مانند زندگی عادی انسان‌ها، در فرآیندها نیز همیشه برخورد، تداخل و رقابت بر سر عوامل مشترک است. هرگز نمی‌توانیم رقابتی را فرض کنیم که هیچ عامل مشترکی در آن نباشد. تمام نزاع‌ها بر سر تصاحب یک «عامل مشترک» است که دو یا چند نفر یا قوم (یا فرآیند یا نخ) به طور همروند یا موازی تلاش می‌کنند تا آن را بدست آورند، یعنی بر روی عامل کاری انجام دهند.

مثال: شرایط رقابتی

دو فرآیند P_1 و P_2 را با کد زیر در نظر بگیرید که در یک سیستم اشتراک زمانی به صورت هم روند اجرا می‌شوند. فرض کنید متغیر a از نوع سراسری و مشترک است و مقدار اولیه آن نیز صفر است، بعد از اجرای کامل دو فرآیند، مقادیر a ، b و c چه خواهد شد؟

کد : $P_۲$	کد : $P_۱$
b=a	a=۱
c=a	

حالت اول:

کد : $P_۲$	کد : $P_۱$
② b=a	① a=۱
③ c=a	

خروجی : a=۱ , b=۱ , c=۱

حالت دوم:

کد : $P_۲$	کد : $P_۱$
① b=a	② a=۱
③ c=a	

خروجی : a=۱ , b=۰ , c=۱

حالت سوم:

کد : $P_۲$	کد : $P_۱$
① b=a	③ a=۱
②c=a	

خروجی : a=۱ , b=۰ , c=۰

توجه :اما مشکل اینجاست که مقدار نهایی متغیرهای a ، b و c، به نحوه‌ی **تعویض متن پردازنده** یا به عبارتی، به ترتیب اجرای دستورالعمل‌ها، بستگی دارد و می‌توانند مقادیر مختلفی را داشته باشند. این پدیده، حاصل رقابت بر سر تصاحب یک عامل مشترک (متغیر مشترک a) است.

مثال: شرایط رقابتی

دو فرآیند هم روند $P_۱$ و $P_۲$ در یک سیستم اشتراک زمانی که از متغیر مشترک سراسری S ، در بخشی از کد خود استفاده می‌کنند در نظر بگیرید، بعد از اجرای کامل دو فرآیند، مقدار نهایی S چه خواهد شد؟ (مقدار اولیه متغیر سراسری S برابر صفر است)

$P_۱$:	$P_۲$:
S=S+۱	S=S-۱

از آنجا که این فرآیندها به زبان اسمبلی یک ماشین فرضی در نظر گرفته می‌شوند، لذا در ادامه دستورات فوق را به صورت سطح غیرانتزاعی‌تر (نمایش جزئیات) و در سطح اسمبلی بازنویسی

می‌کنیم:

P_1:	P_2:
MOVE RGISTER , S	MOVE REGISTER , S
INC REGISTER	DEC REGISTER
MOVE S , REGISTER	MOVE S, REGISTER

حالت اول:

فرض کنید P_1 کامل اجرا شود و سپس P_2 کامل اجرا شود (یا برعکس).

P_1:	P_2:
① REGISTER ← ۰	④ REGISTER ← ۱
② REGISTER ← ۱	⑤ REGISTER ← ۰
③ S ← ۱	⑥ S ← ۰

بنابراین مقدار نهایی متغیر S برابر صفر خواهد بود. (S=۰)

حالت دوم:

فرض کنید ابتدا ترتیب زیر اجرا شود.

P_1:	P_2:
① REGISTER ← ۰	② REGISTER ← ۰
INC REGISTER	③ REGISTER ← - ۱
MOVE S, REGISTER	④ S ← - ۱

سپس پردازنده در اثر تعویض متن به فرآیند P_1 باز گردد.

توجه: هر فرآیند محتویات رجیسترهای خودش را قبل از تعویض متن در PCB ذخیره می‌کند. بنابراین در این لحظه مقدار رجیستر در PCB فرآیند P_1، برابر صفر است.

حال در ادامه دستورات باقی‌مانده فرآیند P_1 اجرا می‌شوند.

P_1:
⋮
⑤ REGISTER ← ۱
⑥ S ← ۱

بنابراین مقدار نهایی متغیر S برابر مثبت یک خواهد بود. (S=+۱)

حالت سوم:

فرض کنید ابتدا ترتیب زیر اجرا شود.

$P_۱$:	$P_۲$:
② REGISTER ← ۰	① REGISTER ← ۰
③ REGISTER ← ۱	DECREGISTER
④ S ← ۱	MOVES, REGISTER

سپس پردازنده در اثر تعویض متن به فرآیند $P_۲$ باز گردد.

توجه : هر فرآیند محتویات رجیسترهای خودش را قبل از تعویض متن در PCB ذخیره می‌کند. بنابراین در این لحظه مقدار رجیستر در PCB فرآیند $P_۲$ ، برابر صفر است.

حال در ادامه دستورات باقی‌مانده فرآیند $P_۲$ اجرا می شوند.

$P_۲$:
⋮
⑤ REGISTER ← - ۱
⑥ S ← - ۱

بنابراین مقدار نهایی متغیر S برابر منفی یک خواهد بود. (S=-۱)

توجه : اما مشکل اینجاست که مقدار نهایی متغیر S ، به نحوه‌ی **تعویض متن پردازنده** یا به عبارتی، به ترتیب اجرای دستورالعمل‌ها، بستگی دارد و می‌تواند مقادیر ۰، ۱ و ۱ـ را داشته باشد. این پدیده، حاصل رقابت بر سر تصاحب یک عامل مشترک (متغیر مشترک S) است.

توجه :وجود پدیده‌ی رقابت در دو مثال قبل در سیستم‌های تک پردازنده‌ای ناشی از وقفه‌ای است که می‌تواند اجرای دستورالعمل‌ها را در هر کجای فرآیند متوقف نماید. (پدیده‌ی تعویض متن). این وضعیت در سیستم‌های چند پردازنده‌ای نیز ممکن است پیش بیاید، به علاوه این که دو یا چند فرآیند می‌توانند به موازات هم اجرا شده و برای دسترسی به یک عامل مشترک در رقابت باشند.

برای کنترل شرایط رقابتی، باید راه حلی ارائه شود که سه شرط زیر را رعایت کند:

۱ ـ شرط انحصار متقابل

برای برقراری شرط انحصار متقابل، عامل مشترک را اسکورت کنید، مانند زمانی که وارد باجه‌ی تلفن همگانی (عامل مشترک) می‌شوید، در را می‌بندید تا مانع ورود شخص دیگری گردید! در عالم انسان‌ها، هیچ دو فردی نباید به طور همزمان وارد عامل مشترک شوند. در عالم فرآیندها نیز هیچ دو فرآیندی نباید به طور همزمان وارد عامل مشترک (ناحیه‌ی بحرانی) شوند. استفاده‌ی همزمان از عامل مشترک معنا ندارد! (اخلاقی نیست) بنابراین باید راهی را پیدا کنیم که از ورود همزمان دو یا چند فرآیند به ناحیه‌ی بحرانی جلوگیری کند. به عبارت دیگر، آنچه که ما به آن نیاز داریم، انحصار متقابل است که در متون فارسی به آن دو به دو ناسازگاری یا مانعةالجمعی نیز گفته می شود، یعنی اگر یکی از فرآیندها

در حال استفاده از حافظه‌ی اشتراکی، فایل اشتراکی و یا هر عامل اشتراکی رقابت‌زاست باید مطمئن باشیم که دیگر فرآیندها، در آن زمان از انجام همان کار محروم می‌باشند. در واقع از بین تمام فرآیندها، در هر لحظه تنها یک فرآیند مجاز است، در عامل مشترک باشد. بدین معنی که اگر فرآیندی در ناحیه‌ی بحرانی است، از ورود فرآیندهای دیگر به همان ناحیه‌ی بحرانی جلوگیری شود و تا خارج شدن فرآیند اول منتظر بمانند، زیرا هیچ دو فرآیندی نباید به طور همزمان وارد ناحیه‌ی بحرانی شوند. به یاد داشته باشید که استفاده‌ی همزمان از عامل مشترک معنا ندارد!

بنابراین برای برقراری شرط انحصار متقابل باید ساختاری را طراحی کنیم که در هر لحظه فقط یک فرآیند مجوز ورود به ناحیه‌ی بحرانی را داشته باشد. لذا هر فرآیند برای ورود به بخش بحرانی‌اش باید اجازه بگیرد. بخشی از کد فرآیند که این اجازه گرفتن را پیاده‌سازی می‌کند، بخش ورودی نام دارد. بخش بحرانی می‌تواند با بخش خروجی دنبال شود. این بخش خروجی کاری می‌کند که فرآیندهای دیگر بتوانند وارد ناحیه‌ی بحرانی‌شان بشوند. بقیه‌ی کد فرآیند را بخش باقی‌مانده می‌نامند. بنابراین ساختار کلی فرآیندها برای برقراری شرط انحصار متقابل به صورت زیر می‌باشد:

```
P (int i) {
   while (TRUE) {
      entry_section (); // تلاش برای کسب اجازه‌ی ورود به ناحیه‌ی بحرانی
      critical_section (); // ناحیه‌ی بحرانی
      exit_section (); // اعلام خروج از ناحیه‌ی بحرانی
      remainder_sectioln (); // ناحیه‌ی باقی‌مانده
      }
   }
```

۲ـ شرط پیشرفت

فرآیندی که داوطلب ورود به ناحیه‌ی بحرانی نیست و نیز در ناحیه‌ی بحرانی قرار ندارد، نباید در رقابت برای ورود سایر فرآیندها به ناحیه‌ی بحرانی شرکت کند، به عبارت دیگر، نباید مانع ورود فرآیندهای دیگر به ناحیه‌ی بحرانی شود. در یک بیان ساده‌تر می‌توان گفت، فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، حق جلوگیری از ورود فرآیندهای دیگر به ناحیه‌ی بحرانی را ندارد. یعنی نباید در تصمیم‌گیری برای ورود فرآیندها به ناحیه‌ی بحرانی شرکت کند.

۳ـ شرط انتظار محدود

فرآیندهایی که نیاز به ورود به ناحیه‌ی بحرانی دارند، باید مدت انتظارشان محدود باشد، یعنی نباید به طور نامحدود در حالت انتظار باقی بمانند.

انتظار نامحدود بر دو دسته می‌باشد: (۱) قحطی، (۲) بن‌بست، بنابراین نباید در شرایط رقابتی بین فرآیندها، قحطی یا بن‌بست رخ دهد.

قحطی (گرسنگی)

در عالم زندگی قحطی زمانی رخ می‌دهد که عده‌ای مدام از منابع مشترک استفاده کنند، و عده‌ای دیگر قادر به استفاده از منابع مشترک نباشند. زیرا دسته‌ی اول از اختصاص منابع به دسته‌ی دوم به طور مداوم و بدون رعایت یک حد بالای مشخص جلوگیری می‌کنند. در عالم فرآیندها نیز هرگاه فرآیندی به مدت نامعلوم و بدون رعایت یک حد بالای مشخص در انتظار گرفتن یک منبع بحرانی یا دسترسی به یک عامل مشترک بماند و فرآیندی دیگر مدام در حال استفاده از منبع بحرانی باشد، در این حالت فرآیند اول دچار قحطی شده است. بنابراین در صورت اقدام یک فرآیند برای ورود به ناحیه‌ی بحرانی، باید محدودیتی برای تعداد دفعاتی که سایر فرآیندها می‌توانند وارد ناحیه‌ی بحرانی شوند، وجود داشته باشد تا قحطی رخ ندهد.

بن‌بست

به وضعیتی که در آن مجموعه‌ای متشکل از دو یا چند فرآیند برای همیشه منتظر یکدیگر بمانند (مسدود) و به عبارت دیگر دچار سیکل انتظار ابدی شوند، بن‌بست گفته می‌شود. به یاد آوردید که شرط انحصار متقابل شرط لازم، شرط انحصاری بودن و شرط نگهداری و انتظار شروط بدیهی و شرط سیکل انتظار (انتظار چرخشی) شرط کافی برای وقوع بن‌بست بود. برای مثال همه فرآیندها ممکن است، منتظر باشند تا فرآیند دیگری از همین مجموعه، یک منبع بحرانی را آزاد کند.

توجه : به تفاوت قحطی و بن‌بست دقت کنید، در قحطی فرآیندی مدام در حال کار و فرآیندی دیگر به مدت نامعلوم در انتظار است. اما در بن‌بست، مجموعه‌ای از فرآیندها در سیکل انتظار ابدی، گرفتار شده‌اند. نه راه پس دارند و نه راه پیش.

توجه : در کنترل شرایط رقابتی، رعایت شرط انحصار متقابل، شرط لازم و رعایت شروط پیشروی و انتظار محدود، شروط کافی برای ارائه‌ی یک راه حل جامع به شمار می‌آیند.

۴ـ همگام‌سازی

گاهی اوقات خواسته ما این است که فرآیندها به یک ترتیب مشخص و از قبل تعیین شده بر روی یک عامل مشترک (داده مشترک) عملیاتی را انجام دهند. واضح است که در این حالت تبادل داده به روش استفاده از حافظه‌ی مشترک است. همچنین از آنجا که پای یک عامل مشترک درمیان است، پس رقابت بر سر تصاحب این عامل مشترک هم در میان است. بنابراین راه حل همگام‌سازی باید به گونه‌ای باشد که فرآیندهای همکار دچار شرایط رقابتی نشوند. به بیان دیگر باید مانع اثر مخرب **تعویض متن**

پردازنده بین فرآیندهای همکار که عامل ایجاد شرایط رقابتی است، شد.
مثلاً اگر اول فرآیند P_1 داده‌ای را باید تولید کند و سپس فرآیند P_2 آن را مصرف کند، فرآیند P_2 باید منتظر باشد تا فرآیند P_1 ، داده مورد نیازش را آماده کند و بعد شروع به مصرف نماید.

مثال: همگام‌سازی دو فرآیند

دو فرآیند P_1 و P_2 را در نظر بگیرید که در یک سیستم تک پردازنده اشتراک زمانی به صورت هم روند اجرا می‌شوند. فرض کنید فرآیند P_1 در بخشی از کد خود مقدار متغیر x را می‌خواند و فرآیند P_2 نیز در بخشی از برنامه‌اش باید مقدار متغیر x خوانده شده توسط P_1 را چاپ کند. هدف مسأله، نوشتن این دو برنامه به صورتی است که P_2 در چاپ متغیر x بر P_1 پیشی نگیرد و اگر زودتر به بخش چاپ متغیر x رسید و هنوز مقدار متغیر x خوانده نشده بود، صبر کند تا P_1 متغیر x را مقداردهی کند. بنابراین، باید یک مسأله‌ی همگام‌سازی حل شود. واضح است که در این حالت تبادل داده به روش استفاده از حافظه‌ی مشترک است. یکی از راه حل‌های مناسب برای همگام‌سازی فرآیندهای همکار استفاده از سمافور می‌باشد که جلوتر شرح خواهیم داد.

راه حل‌های کنترل شرایط رقابتی

۱ـ راه حل‌های نرم‌افزاری (بر عهده‌ی برنامه‌نویس)
۲ـ راه حل‌های سخت‌افزاری
۳ـ راه حل‌های سیستم عامل (سمافور)
۴ـ راه حل‌های زبان‌های برنامه‌سازی (مانیتور)

توجه : راه حل‌های فوق، در سیستم‌های تک پردازنده و چند پردازنده با حافظه‌ی اشتراکی قابل استفاده‌اند.

راه حل‌های نرم‌افزاری

در این دسته راه حل‌ها، مسئولیت برقراری شرط انحصار متقابل بر عهده‌ی خود فرآیندها (کد نوشته شده توسط برنامه نویس) است و هیچ حمایتی از زبان برنامه‌سازی و سیستم عامل وجود ندارد.

۱ ـ راه حل متغیر قفل

یکی از راه حل‌های نرم‌افزاری کنترل شرایط رقابتی، استفاده از متغیر قفل می‌باشد. برای تجسم بهتر، یک باجه‌ی تلفن همگانی را به عنوان یک عامل مشترک در نظر بگیرید که قصد داریم شرط انحصار متقابل را برای آن برقرار کنیم. برای این کار، یک تابلوی اعلام وضعیت، در ورودی درب باجه نصب می‌کنیم. اگر تابلو عدد صفر را نمایش داد، بدین معنی است که فردی داخل نیست (باجه خالی است) و اگر تابلو عدد یک را نمایش داد، بدین معنی است که فردی داخل است (باجه پُر است).

اما در این راه حل ساده، شرط انحصار متقابل برقرار نیست، حالتی را در نظر بگیرید که تابلو به نشانه‌ی خالی بودن باجه، عدد صفر را نمایش می‌دهد، بنابراین ممکن است دو فرد، به شکل همزمان تابلوی وضعیت را به نشانه‌ی خالی بودن باجه مشاهده کنند و هر دو با هم اقدام به ورود به باجه تلفن (عامل مشترک) نمایند. بنابراین در این شرایط، شرط انحصار متقابل به عنوان شرط لازم شروط کنترل شرایط رقابتی نقض شده است.

شرح الگوریتم

فرض کنید یک متغیر قفل یکتای مشترک، با مقدار اولیه صفر موجود است. هنگامی که فرآیندی می‌خواهد وارد ناحیه‌ی بحرانی خود شود، ابتدا قفل را تست می‌کند، اگر قفل صفر باشد آن را برابر یک قرار می‌دهد و وارد ناحیه‌ی بحرانی خود می‌شود، ولی اگر قفل برابر یک بود، باید منتظر بماند تا قفل برابر صفر شود. بنابراین صفر به این معنی است که هیچ فرآیندی در ناحیه‌ی بحرانی قرار ندارد و یک به معنی این است که یک فرآیند در ناحیه‌ی بحرانی‌اش قرار دارد.

ساختار کلی این راه حل به صورت زیر می‌باشد:

P_0

```
P (int i) {
   while (TRUE) {
      while (lock ==1); /*loop*/
      lock=1;
      critical_section();
      lock=0;
      reminder_section();
      }
   }
```

P_1

```
P (int i) {
   while (TRUE) {
      while (lock ==1); /*loop*/
      lock=1;
      critical_section();
      lock=0;
      reminder_section();
      }
   }
```

توجه : قرار دادن کاراکتر؛ در انتهای حلقه while، سبب می‌شود حلقه تا زمانی که شرط برقرار است، در خود بچرخد.

اما این راه حل، با وجود اینکه خیلی ساده به نظر می‌رسد، ولی شرط انحصار متقابل را رعایت نمی‌کند. سناریوی زیر را در نظر بگیرید:

فرض کنید فرآیند P_0 ، در حلقه، متغیر قفل را می‌خواند و آن را برابر صفر می‌بیند، بنابراین شرط حلقه برقرار نیست و کنترل برنامه به خط بعد می‌رود، ولی قبل از آنکه بتواند مقدار یک را درون متغیر قفل قرار دهد، پردازنده به فرآیند P_1 تعویض متن می‌کند. حال فرآیند P_1 نیز متغیر قفل را برابر صفر

می‌بیند، بنابراین شرط حلقه برقرار نیست و کنترل برنامه به خط بعد می‌رود، در نتیجه مقدار متغیر قفل را برابر یک قرار می‌دهد و در ادامه وارد ناحیه‌ی بحرانی می‌شود. حال اگر دوباره در همین لحظه پردازنده به فرآیند P_0 تعویض متن کند، فرآیند P_0 نیز قفل را برابر یک قرار می‌دهد و وارد ناحیه‌ی بحرانی‌اش می‌شود، یعنی هر دو فرآیند P_0 و P_1 همزمان در عامل مشترک (ناحیه‌ی بحرانی) قرار دارند. این یعنی نقض شرط انحصار متقابل در شرایط رقابتی ما بین فرآیندها. پس شرط انحصار متقابل به عنوان شرط لازم در کنترل شرایط رقابتی برقرار نیست. بنابراین این راه حل مناسب نیست.

۲ ـ راه حل تناوب قطعی

یکی دیگر از راه حل‌های نرم‌افزاری کنترل شرایط رقابتی، استفاده از تناوب قطعی می‌باشد. برای تجسم بهتر، یک باجه‌ی تلفن را به عنوان یک عامل مشترک در نظر بگیرید که قصد داریم شرط انحصار متقابل را برای آن برقرار کنیم. برای این کار، این بار یک تابلوی اعلام نوبت برخلاف راه قبل که یک تابلوی اعلام وضعیت ناحیه‌ی بحرانی بود، در ورودی، درب باجه نصب می‌کنیم. اگر تابلو عدد صفر را نمایش داد، بدین معنی است که نوبت فرآیند P_0 است و اگر تابلو عدد یک را نمایش داد، بدین معنی است که نوبت فرآیند P_1 است. اما در این راه حل ساده، این بار شرط پیشروی برقرار نیست.

در شرط پیشرفت قرارمان این بود که اگر فردی داخل ناحیه‌ی باقی‌مانده بود، از ورود فرد دیگری به ناحیه‌ی بحرانی جلوگیری نکند، این همان تعریف شرط پیشرفت بود، اما این راه حل بر سر قرار پیشرفت نمانده است!

حالتی را در نظر بگیرید که تابلوی نوبت، عدد صفر را نشان می‌دهد، بنابراین در حال حاضر نوبت فرد P_0 برای ورود به باجه‌ی تلفن (عامل مشترک) است. پس P_0 وارد باجه‌ی تلفن می‌شود، و فرد دیگری حق ورود به باجه‌ی تلفن را ندارد، زیرا نوبتش نیست! این یعنی برقراری شرط انحصار متقابل. اکنون فرد P_0 داخل باجه قرار دارد و پس از آنکه تلفنش تمام شد، تابلوی نوبت را به یک مقداردهی می‌کند، از باجه خارج می‌شود و می‌رود دنبال کار کپی تا اسنادی را کپی نماید و مدت زیادی را مشغول این کار می‌ماند، در واقع P_0 در حال حاضر، در ناحیه‌ی باقی‌مانده کارهای خود قرار دارد.

در این لحظه فرد P_1 از راه می‌رسد، تابلوی نوبت را می‌بیند که مقدار آن برابر یک است، پس نوبتی هم که باشد این بار نوبت فرد P_1 است. بنابراین P_1 وارد باجه‌ی تلفن می‌شود، و فرد دیگری حق ورود به باجه‌ی تلفن را ندارد، زیرا نوبتش نیست! این یعنی برقراری انحصار متقابل. اکنون فرد P_1 داخل باجه قرار دارد و پس از آنکه تلفنش تمام شد، تابلوی نوبت را به صفر مقداردهی می‌کند و از باجه خارج می‌شود. حال دوباره نوبت فرد P_0 است. اما همانطور که گفتیم فرد P_0 در حال کپی اسنادی می‌باشد، همچنان هم کار کپی‌اش ادامه دارد...

یعنی فرد P_0 همچنان در ناحیه‌ی باقی‌مانده قرار دارد، در این لحظه فرد P_1 ناگهان تصمیم می‌گیرد تا

دوباره تلفن بزند، اما نمی‌تواند داخل باجه‌ی تلفن شود، چون نوبتش نیست، نوبتی هم که باشد این بار نوبت فرد P_0 است تا وارد باجه شود، پس فرد P_1 باید صبر پیشه کند، تا فرد P_0 کپی‌اش تمام شود و برود داخل باجه، تلفنش را بزند، کارش که تمام شد، تابلوی نوبت را به یک مقداردهی کند تا بالاخره نوبت فرد P_1 شود. نه، قرار این نبود، قرار این نبود که فردی که داخل ناحیه‌ی باقی‌مانده قرار دارد، از ورود فرد دیگری به باجه‌ی تلفن جلوگیری کند. دیدید که فرد P_0 با اینکه در ناحیه‌ی باقی‌مانده قرار داشت چگونه مانع ورود فرد P_1 به باجه‌ی تلفن شد. پس این راه حل هم مناسب نیست، به کار موقعیتی می‌آید، که مدام افراد نبوتی، تلفن می‌زنند و نوبت را سریع به طرف دیگر می‌سپارند، یعنی تناوب قطعی دارند.

شرح الگوریتم

P_0	P_1
P(int i){	P(int i){
while (TRUE)	while (TRUE)
{	{
① while (turn!=0); /*loop*/	while (turn!=1); /*loop*/
②/ critical_section ();	critical_section ();
③ turn=1;	turn=0;
④remainder_section ( );	remainder_section ( );
}	}
}	}

توجه : در این الگوریتم امکان ندارد دو فرآیند با هم وارد ناحیه‌ی بحرانی شوند، علت این امر به خاصیت الا کلنگی روش نوبت‌بندی مربوط می‌شود. بنابراین شرط انحصار متقابل برقرار است.

در این الگوریتم یک متغیر سراسری و مشترک به نام turn تعریف می‌شود که این متغیر، نوبت فرآیندها را برای ورود به ناحیه‌ی بحرانی نگه می‌دارد. زمانی که متغیر turn برابر صفر باشد، با اجرای فرآیند P_0، این فرآیند متغیر turn را بررسی و مقدار آن را مساوی صفر می‌بیند (خط ①)، بنابراین از حلقه‌ی while عبور کرده (چون شرط حلقه برقرار نیست) و وارد ناحیه‌ی بحرانی می‌گردد. حال اگر فرآیند P_1 نیز بخواهد به ناحیه‌ی بحرانی دسترسی داشته باشد با چک کردن مقدار این متغیر (در خط ①) منتظر مانده و مجوز ورود به ناحیه‌ی بحرانی به آن داده نمی‌شود، چون شرط حلقه در آن برقرار بوده و در اجرای دستور while می‌چرخد. بنابراین در داخل یک حلقه‌ی انتظار مشغول، بیکار می‌ماند و مرتباً turn را تست می‌کند تا ببیند چه موقع مقدار turn برابر یک می‌شود. زمانی که فرآیند P_0 از ناحیه‌ی بحرانی خارج می‌شود. مقدار turn را برابر یک قرار می‌دهد تا فرآیند P_1 بتواند وارد ناحیه‌ی بحرانی

شود.

مزایا

رعایت شرط انحصار متقابل

فرآیند P_0 وقتی وارد می‌شود که مقدار متغیر turn برابر صفر باشد و فرآیند P_1 وقتی وارد می‌شود که مقدار متغیر turn برابر یک باشد. بنابراین امکان ندارد فرآیندها با هم وارد ناحیه‌ی بحرانی شوند، زیرا مقدار متغیر turn یا صفر است و یا یک.

فرض کنید فرآیند P_0 در ناحیه‌ی بحرانی قرار دارد. یعنی مقدار متغیر turn برابر صفر است. از طرفی فرآیند P_1 تلاش می‌کند که به ناحیه‌ی بحرانی وارد شود. شرط ورود این فرآیند، یک شدن مقدار متغیر turn است، یعنی باید مقدار متغیر turn برابر یک شود، که این امر تنها در صورت خروج فرآیند P_0 از ناحیه‌ی بحرانی به وقوع می‌پیوندد. چنانچه فرآیند P_1 نیز در ناحیه‌ی بحرانی باشد، همین وضعیت برای فرآیند P_0 تکرار می‌شود. در نتیجه دو فرآیند هیچ‌گاه همزمان در ناحیه‌ی بحرانی قرار نمی‌گیرند، بنابراین شرط انحصار متقابل همواره برقرار است.

رعایت شرط انتظار محدود (عدم بن‌بست و قحطی)

در این الگوریتم، هر فرآیند پس از یک نوبت انتظار، می‌تواند وارد ناحیه‌ی بحرانی شود. در یک سناریو از اجرا، فرض کنید فرآیند P_1 منتظر ورود به ناحیه‌ی بحرانی است که چون اکنون P_0 در ناحیه‌ی بحرانی قرار دارد به آن اجازه‌ی ورود داده نمی‌شود. حال اگر فرایند P_0 از قسمت ناحیه‌ی بحرانی بگذرد و وارد بخش اعلام خروج شود، یعنی مقدار متغیر turn را یک کند و سپس بخواهد بدون توجه به درخواست فرآیند دیگر، مجدداً وارد ناحیه‌ی بحرانی گردد، به آن اجازه داده نمی‌شود و باید صبر کند تا P_1 یک بار اجرا شده و در انتهای کار خودش مقدار متغیر turn را صفر کند و مجوز ورود مجدد را به فرآیند P_0 بدهد. پس در این الگوریتم گرسنگی وجود ندارد. زیرا فرآیندها به صورت یک در میان و نوبتی وارد ناحیه‌ی بحرانی می‌شوند و نه تصادفی.

گرسنگی یعنی اینکه که یک فرآیند مدام کار کند و از کار کردن فرآیندی دیگر جلوگیری کند. در اینجا گرسنگی بدین معنی که یک فرآیند مدام داخل ناحیه‌ی بحرانی برود و از ورود فرآیندی دیگر جلوگیری کند وجود ندارد. از آن جا که در اجرای همروند و موازی این دو فرآیند، بن‌بست نیز وجود ندارد لذا شرط انتظار محدود برقرار است.

چگونگی امکان وقوع بن‌بست

هرگاه دو فرآیند متقاضی ورود به ناحیه‌ی بحرانی به طور همزمان تا ابد منتظر ورود به ناحیه‌ی بحرانی باشند، در این شرایط هر دو فرآیند مسدود و به خواب رفته‌اند، در این حالت بن‌بست رخ داده است.

معایب

عدم رعایت شرط پیشرفت

فرض کنید فرآیند P_1 یک باقی‌مانده‌ی سنگین دارد (مانند رسم یک نمودار طولانی) و مشغول انجام این کار در ناحیه‌ی باقی‌مانده است (خط ④). از طرفی P_0 در این لحظه از ناحیه‌ی بحرانی بیرون آمده و مقدار متغیر turn را در ناحیه خروج برابر یک قرار داده است و پس از مدت کمی، مجدداً درخواست ورود به ناحیه‌ی بحرانی را دارد. اما از آنجا که فرآیند P_1 هنوز در ناحیه‌ی باقی‌مانده خود به سرمی‌برد و نتوانسته است تاکنون با دستیابی مجدد به ناحیه‌ی بحرانی، مقدار متغیر turn را صفر کند (یعنی اجرای خط ③). بنابراین فرآیند P_0 مجبور است منتظر انجام بخش باقی‌مانده‌ی اجرای فرآیند مقدار P_1 که در ناحیه‌ی بحرانی قرار ندارد، شود. پس شرط پیشرفت در این الگوریتم نقض می‌گردد.

وجود پدیده‌ی گلوگاه

در این الگوریتم فرآیندها برای دستیابی به ناحیه‌ی بحرانی باید به صورت یک در میان عمل کنند، بنابراین سرعت اجرای عملیات توسط فرآیند کندتر هدایت می‌شود (گلوگاه).

مسأله انتظار مشغول

در این راه حل مشکل انتظار مشغول[1] وجود دارد، زیرا در زمانی که مثلاً فرآیند P_0 در ناحیه‌ی بحرانی قرار دارد، فرآیند P_1 زمان پردازنده را در چک کردن مقدار متغیر turn در خط ۱ ، در هر بار تعویض متن به فرآیند P_1 مصرف می‌کند و عملاً زمان پردازنده بیهوده هدر می‌رود.

توجه :روش نوبت گرفتن در مواقعی که یکی از فرآیندها خیلی کندتر است، ایده‌ی خوبی نیست.

توجه : در برخی از کتب مرجع و غیر مرجع، قبل از بیان راه حل پترسون، راه حل‌های ابتدایی‌تری موسوم به تلاش دوم، سوم و ... مطرح شده است، اما ما صلاح دیدیم ابتدا راه حل پترسون را معرفی کنیم و سپس قطعاتی از راه حل پترسون را برداریم تا به همان راه حل‌های ابتدایی قبل از پترسون برسیم، تا ارزش و قدرت راه حل پترسون هرچه بیشتر و بیشتر پر رنگ‌تر و با شکوه‌تر جلوه کند.

راه حل پترسون

این راه حل، ساده‌ترین و کوتاه‌ترین راه حل نرم‌افزاری است. در این راه حل از یک متغیر نوبت و یک تابلوی وضعیت دو یا چند حالته برای فرآیندها استفاده می‌گردد.

متغیر نوبت :برای نوبت گرفتن فرآیندها و ضایع نشدن حق یک فرآیند در ورود به ناحیه‌ی بحرانی مورد استفاده قرار می‌گیرد. بنابراین شرط انتظار محدود برقرار می‌شود.

1- Busy waiting

تابلوی وضعیت فرآیند: وضعیت فعلی یک فرآیند را برای فرآیند رقیب به نمایش می‌گذارد. تا اگر یکی از فرآیندها در ناحیه‌ی بحرانی نبود و یا قصد ورود به ناحیه‌ی بحرانی را نیز نداشت، یعنی در ناحیه‌ی باقی‌مانده قرار داشت، فرآیند دیگری بتواند به ناحیه‌ی بحرانی دسترسی داشته باشد. بنابراین شرط پیشروی برقرار می‌شود.

توجه : تابلوی وضعیت فرآیندها، سبب مستقل شدن فرآیندها از تناوب قطعی و خاصیت الاکلنگی می‌شود. این تابلو، دو وضعیت مختلف را برای یک فرآیند به نمایش می‌گذارد:

Flag[i] = FALSE : یعنی فرآیند مورد نظر در داخل ناحیه‌ی بحرانی قرار ندارد.

Flag[i] = TRUE : یعنی فرآیند مورد نظر علاقه‌مند است تا در صورت خالی بودن ناحیه‌ی بحرانی، وارد ناحیه‌ی بحرانی گردد. به عبارت دیگر علاقه‌مندی یک فرآیند برای ورود به ناحیه‌ی بحرانی را نشان می‌دهد.

شرط ورود به ناحیه‌ی بحرانی یک فرآیند در راه حل پترسون

شروط لازم

- فرآیند، علاقه‌مند به ورود به ناحیه‌ی بحرانی باشد، Flag [فرآیند علاقه‌مند] =TRUE
- فرآیند رقیب در ناحیه‌ی بحرانی نباشد، Flag [فرآیند رقیب] =FALSE

شرط کافی

- فرآیند مورد نظر، متغیر نوبت turn را زودتر مقداردهی کند. در واقع سرنوشت نهایی ورود فرآیندها به ناحیه‌ی بحرانی به متغیر نوبت turn گره خورده است. در یک قاعده‌ی کلی، در رقابت بر سر تصاحب عامل مشترک (ناحیه‌ی بحرانی) هر فرآیندی که علاقه‌مند به ورود به ناحیه‌ی بحرانی باشد (Flag [i]=TRUE) و سپس زودتر متغیر نوبت (turn=i) را مقداردهی کند، تحت هر شرایطی، تأکید می‌کنیم، تحت هر شرایطی زودتر وارد ناحیه‌ی بحرانی می‌شود و فرآیندی که دیرتر متغیر نوبت turn را مقدار دهی کند، پس از مقداردهی متغیر نوبت (turn=i)، باید بنشیند و در یک حلقه‌ی انتظار، صبر پیشه کند و مقدار متغیر نوبت turn را نگه‌داری کند تا فرآیند رقیب از ناحیه‌ی بحرانی خارج شود یعنی ناحیه‌ی بحرانی خالی گردد.

ساختار کلی این راه حل به صورت زیر می‌باشد:

```
Boolean flag[2]= {FALSE, FALSE};
int turn;
```

P_0	P_1
P_0 (void) {	P_1 (void) {
while (TRUE)	while (TRUE)
{	{
①flag [0]= TRUE;	flag [1]= TRUE;
②turn = 0;	turn = 1;
③while (flag[1] && turn == 0); /*loop*/	while(flag[0] &&turn== 1); /*loop*/
④critical_section ();	critical_section ();
⑤flag [0] = FALSE;	flag [1] = FALSE;
remainder_section ();	remainder_section ();
}	}
}	}

مزایا

رعایت شرط انحصار متقابل

حالت اول (ورود غیرهمزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی باشد. یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی نباشد. پس در ابتدا Flag[0]=FALSE, Flag [1] = FALSE می‌باشد. فرض کنید فرآیند P_0 با مقداردهی Flag [0] = TRUE علاقه‌مندی خود را برای ورود به ناحیه‌ی بحرانی اعلام کند و در ادامه با مقداردهی متغیر turn برابر با صفر (turn=۰) به شکل یکه و تنها (اولی و آخری خودش است) پس از عبور از خط ③ به دلیل عدم برقراری شرط حلقه while (flag[1] && turn==0); وارد ناحیه‌ی بحرانی گردد.

FALSE TRUE

توجه کنید شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار بود. حال اگر پردازنده به فرآیند P_1 تعویض متن انجام دهد، فرآیند P_1 در حلقه‌ی انتظار مشغول خواهد ماند.

while (flag[0] & & turn = = 1);

TRUE TRUE

تا زمانی که فرآیند P_0 از ناحیه‌ی بحرانی خود خارج شود و Flag [0] = FALSE شود. توجه کنید که

شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_1 برقرار نبود. زیرا ناحیه‌ی بحرانی پُر است. بنابراین شرط انحصار متقابل برقرار است.

حالت دوم (ورود تقریباً همزمان فرآیندها)

زمانی که هر دو فرآیند به طور تقریباً همزمان، قصد ورود به ناحیه‌ی بحرانی را دارند. بنابراین تابلوی وضعیت فرآیندها هر دو Flag [0] = TRUE و Flag [1] = TRUE می‌شود، زیرا هر دو فرآیند به شکل تقریباً عمزمان علاقه‌مند به ورود به ناحیه‌ی بحرانی هستند. اما متغیر نوبت turn نمی‌تواند در یک زمان هم صفر و هم یک باشد. زیرا پس از آنکه هر دو فرآیند، شماره‌ی فرآیند خود را در متغیر نوبت turn ذخیره نمودند. فرآیندی که دیرتر شماره‌اش را ذخیره کند، فرآیندی است که شماره‌اش در متغیر نوبت turn باقی می‌ماند و دیگری اثرش در متغیر نوبت turn از بین می‌رود. در واقع سرنوشت ورود فرآیندها به ناحیه‌ی بحرانی به متغیر نوبت turn گره خورده است (شرط کافی)، بنابراین فرآیندی که دیرتر متغیر نوبت turn را مقداردهی کرده است، باید صبر پیشه کند و متغیر نوبت turn را نگه‌داری کند و در یک حلقه‌ی انتظار بچرخد.

فرض کنید، فرآیند P_1 دیرتر اقدام به ورود به ناحیه‌ی بحرانی کند، بنابراین مقدار متغیر نوبت turn برابر با یک خواهد بود. $(turn = 1)$ وقتی که هر دو فرآیند به دستور while می‌رسند، خط ③ برای فرآیند P_0 برقرار نیست.

```
while (flag [1] && turn == 0);
       TRUE          FALSE
```

بنابراین در حلقه نمی‌چرخد، پس فرآیند P_0 وارد ناحیه‌ی بحرانی می‌شود، اما خط ③ برای فرآیند P_1 برقرار است.

```
while (flag [0] && turn == 1);
       TRUE          TRUE
```

بنابراین در حلقه می‌چرخد و وارد ناحیه‌ی بحرانی نمی‌شود و می‌نشیند و صبر پیشه می‌کند و متغیر نوبت turn را نگه‌داری می‌کند. تا زمانی که فرآیند P_0 از ناحیه‌ی بحرانی خود خارج گردد و Flag [0] = FALSE شود. بنابراین شرط انحصار متقابل در این حالت هم برقرار است.

رعایت شرط پیشرفت

شرط پیشرفت می‌گفت، فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، در تصمیم‌گیری برای ورود فرآیندهای دیگر به ناحیه‌ی بحرانی شرکت نکند. فرض کنید فرآیند P_1 در ناحیه‌ی باقی‌مانده‌ی سنگین خود قرار دارد (در حال حاضر در ناحیه‌ی بحرانی قرار ندارد و قصد ورود به ناحیه‌ی بحرانی را هم

ندارد)، مانند رسم یک نمودار طولانی، (بعد از خط ⑤) و در این زمان P_0 از ناحیه‌ی بحرانی خارج گردد و مجدداً درخواست این ناحیه را داشته باشد، اما فرآیند P_1 همچنان در ناحیه‌ی باقی‌مانده‌ی سنگین خود قرار دارد. از آنجا که فرآیند P_1 پس از خروج از ناحیه‌ی بحرانی و قبل از ورود به ناحیه‌ی باقی‌مانده‌ی سنگین خود مقدار Flag [1] را برابر FALSE قرار داده است (در خط ⑤)، پس باز هم فرآیند P_0 بدون وجود مانعی، وارد ناحیه‌ی بحرانی می‌شود. فرآیند P_0 داخل ناحیه‌ی بحرانی می‌گردد، در حالی که فرآیند P_1 در ناحیه‌ی باقی‌مانده قرار دارد، این یعنی پیشرفت. بنابراین شرط پیشرفت برقرار است.

توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار است.

while (flag[1] && turn == 0);

FALSE TRUE

رعایت شرط انتظار محدود

گرسنگی ندارد: گرسنگی یعنی اینکه یک فرآیند مدام کار کند و از کار کردن فرآیندی دیگر جلوگیری کند. در اینجا گرسنگی بدین معنی که یک فرآیند مدام داخل ناحیه‌ی بحرانی برود و از ورود فرآیندی دیگر به ناحیه‌ی بحرانی جلوگیری کند، وجود ندارد. سرنوشت ورود فرآیندها به متغیر نوبت turn گره خورده است، هر فرآیندی که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی، زودتر برایش برقرار باشد، زودتر وارد ناحیه‌ی بحرانی می‌گردد. بنابراین فرآیندی گرسنه نمی‌ماند.

فرض کنید فرآیند P_0 از ناحیه‌ی بحرانی خارج گردد و مقدار Flag [0] را برابر FALSE قرار دهد و برای اینکه فرآیند P_1 را گرسنه نگه دارد، مجدداً قصد ورود به ناحیه‌ی بحرانی را داشته باشد، در حالی که فرآیند P_1 در زمانی که فرآیند P_0 داخل ناحیه‌ی بحرانی بود، منتظر خالی شدن ناحیه‌ی بحرانی بود و نوبت گرفته بود، فرآیند P_0 می‌خواهد، فرآیند P_1 را گرسنه نگه دارد ولی نمی‌تواند، زیرا شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرایند P_0 برقرار نیست.

while (flag [0] & & turn == 0);

TRUE TRUE

زیرا P_1 زودتر از P_0 نوبت گرفته است و باعث می‌شود مقدار متغیر نوبت turn برابر صفر (turn = ۰) باشد. زیرا مقدار صفر مربوط به فرآیند P_0 که دیرتر آمده است در متغیر نوبت turn قرار می‌گیرد. بنابراین فرآیند P_1 وارد ناحیه‌ی بحرانی می‌گردد، زیرا شروط لازم و کافی برای فرآیند P_1 برقرار است.

while (flag [0] && turn = = 1);

FALSE FALSE

لذا این بار فرآیند P_0 باید بنشیند و صبر پیشه کند و مقدار متغیر نوبت turn را نگه‌داری کند تا فرآیند P_1 از ناحیه‌ی بحرانی خارج شود یعنی ناحیه‌ی بحرانی خالی گردد.

بن‌بست ندارد

زیرا پس از اجرای موازی یا همروند خطوط ① و ② ، با توجه به مقدار متغیر نوبت turn، در نهایت یکی از فرآیندها به طور قطع بسته به اینکه چه کسی متغیر turn را زودتر مقداردهی کرده است، وارد ناحیه‌ی بحرانی می‌گردد و هیچگاه دو فرآیند در خط ③ به طور همزمان مسدود نمی‌شوند، در نتیجه بن‌بست ندارد. بنابراین شرط انتظار محدود برقرار است.

معایب

۱ـ عدم رعایت شرط انتظار مشغول

در این راه حل مشکل انتظار مشغول وجود دارد، زیرا زمانی که مثلاً فرآیند P_0 در ناحیه‌ی بحرانی قرار دارد، فرآیند P_1 در یک حلقه‌ی انتظار زمان پردازنده را در بررسی برقراری شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی در خط ③ به هدر می‌دهد.

ناپترسون‌ها

در ادامه با تغییراتی در راه حل پترسون، راه حل‌هایی را ایجاد می‌کنیم که هیچ یک راه حل جامعی نخواهند بود، اما سبب می‌شود تا ارزش راه حل پترسون بیشتر شناخته شود.

ناپترسون اول

حذف متغیر نوبت turn از راه حل پترسون. این راه حل، به سومین تلاش نیز موسوم است.

در این راه حل تنها از یک تابلوی وضعیت دو یا چند حالته برای فرآیندها استفاده می‌گردد.

تابلوی وضعیت فرآیند: وضعیت فعلی یک فرآیند را برای فرآیند رقیب به نمایش می‌گذارد. تا اگر یکی از فرآیندها در ناحیه‌ی بحرانی نبود و یا قصد ورود به ناحیه‌ی بحرانی را نیز نداشت، یعنی در ناحیه‌ی باقی‌مانده قرار داشت، فرآیند دیگری بتواند به ناحیه‌ی بحرانی دسترسی داشته باشد. بنابراین شرط پیشروی برقرار است.

توجه : تابلوی وضعیت فرآیندها، سبب مستقل شدن فرآیندها از تناوب قطعی و خاصیت الاکلنگی می‌شود. این تابلو، دو وضعیت مختلف را برای یک فرآیند به نمایش می‌گذارد:

Flag [i] = FALSE: یعنی فرآیند مورد نظر در داخل ناحیه‌ی بحرانی قرار ندارد.

Flag [i] = TRUE: یعنی فرآیند مورد نظر علاقه‌مند است تا در صورت خالی بودن ناحیه‌ی بحرانی، وارد ناحیه‌ی بحرانی گردد. به عبارت دیگر علاقه‌مندی یک فرآیند برای ورود به ناحیه‌ی بحرانی را نشان می‌دهد.

شرط ورود به ناحیه‌ی بحرانی یک فرآیند در این راه حل

شرط لازم

• فرآیند، علاقه‌مند به ورود به ناحیه‌ی بحرانی باشد، Flag [فرآیند علاقه‌مند] = TRUE

شرط کافی

• فرآیند رقیب در ناحیه‌ی بحرانی نباشد، Flag [فرآیند رقیب] = FALSE

ساختار کلی این راه حل به صورت زیر می‌باشد:

```
Boolean flag[2]= {FALSE, FALSE};
```

P۰

```
P۱(void){
   while (TRUE);
   {
   ①flag [0]= TRUE;
   ②while (flag [1]); /*loop*/
   ③critical_section ( );
   ④flag [0] = false;
   remainder_section ( );
   }
}
```

P۱

```
P۱(void){
while (TRUE);
{
flag [1]= TRUE;
while (flag [0]); /*loop*/
 critical_section ( );
flag [1] = false;
remainder_section ();
}
}
```

توجه : در این راه حل ابتدا فرآیند، علاقه‌مندی خود مبنی بر ورود به ناحیه‌ی بحرانی را با مقداردهی flag[i]=TRUE اعلام می‌کند، سپس وضعیت فرآیند مقابل، بررسی می‌شود.

مزایا

رعایت شرط انحصار متقابل

حالت اول (ورود غیرهمزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی باشد، یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی نباشد. پس در ابتدا flag [1] = FALSE و flag [0] = FALSE می‌باشد. فرض کنید، فرآیند P_0 با مقداردهی flag[0] = TRUE علاقه‌مندی خود را برای ورود به ناحیه‌ی بحرانی اعلام کند و در ادامه پس از عبور از خط ② به دلیل عدم برقراری شرط حلقه while (flag[1]); وارد ناحیه‌ی بحرانی گردد.

FALSE

توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار بود. حال اگر

پردازنده به فرآیند P_1 تعویض متن انجام دهد، فرآیند P_1 در حلقه‌ی انتظار مشغول خواهد ماند.

while (flag [0]);

TRUE

تا زمانی که فرآیند P_0 از ناحیه‌ی بحرانی خود خارج شود و flag [0] = FALSE شود. توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرایند P_1 برقرار نبود. زیرا ناحیه‌ی بحرانی پُر است. بنابراین شرط انحصار در حالت اول برقرار است.

حالت دوم (ورود تقریباً همزمان فرآیندها)

زمانی که هر دو فرآیند به طور تقریباً همزمان، قصد ورود به ناحیه‌ی بحرانی را دارند. بنابراین تابلوی وضعیت فرآیندها هر دو برابر Flag [0] = TRUE و Flag [1] = TRUE می‌شود، زیرا هر دو فرآیند به شکل تقریباً همزمان علاقه‌مند به ورود به ناحیه‌ی بحرانی هستند.

وقتی که هر دو فرآیند به طور تقریباً همزمان به دستور while می‌رسند، شرط خط ②

برای فرآیند P_0

while (flag [0]);

TRUE

برای فرآیند P_1

while (flag [1]);

TRUE

برقرار است. بنابراین هر دو فرآیند P_0 و P_1 در حلقه دور می‌زنند و وارد ناحیه‌ی بحرانی نمی‌شوند. بنابراین شرط انحصار متقابل برقرار است.

اما هر دو فرآیند، تا ابد گرفتار حلقه‌ی انتظار مشغول شده‌اند، بنابراین بن‌بست رخ داده است و شرط انتظار محدود برقرار نیست.

رعایت شرط پیشرفت

شرط پیشرفت می‌گفت، فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، در تصمیم‌گیری برای ورود فرآیندهای دیگر به ناحیه‌ی بحرانی شرکت نکند.

فرض کنید فرآیند P_1 در ناحیه‌ی باقی‌مانده‌ی سنگین خود قرار دارد (در حال حاضر در ناحیه‌ی بحرانی قرار ندارد و قصد ورود به ناحیه‌ی بحرانی را هم ندارد)، مانند رسم یک نمودار طولانی، (بعد از خط ④) و در این زمان P_0 از ناحیه‌ی بحرانی خود خارج گردد و مجدداً درخواست این ناحیه را داشته باشد، اما فرایند P_1 همچنان در ناحیه‌ی باقی‌مانده سنگین خود قرار دارد. از آنجا که فرآیند P_1 پس از خروج از ناحیه‌ی بحرانی و قبل از ورود به ناحیه‌ی باقی‌مانده‌ی سنگین خود، مقدار Flag [1] را

برابر FALSE قرار داده است (در خط ④)، پس باز هم فرآیند P_0 بدون وجود مانعی، وارد ناحیه‌ی بحرانی می‌شود. فرآیند P_0 داخل ناحیه‌ی بحرانی می‌گردد، در حالی که فرآیند P_1 در ناحیه‌ی باقی‌مانده قرار دارد، این یعنی پیشرفت. بنابراین شرط پیشرفت برقرار است.
توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار است.

while (flag [1]);
FALSE

معایب

عدم رعایت شرط انتظار محدود

بن‌بست دارد: زیرا در ورود تقریباً همزمان فرآیندها، بن‌بست وجود داشت.

ناپترسون دوم

حذف متغیر نوبت turn و جابه‌جا کردن خطوط ① و ③ از راه حل پترسون. این راه حل به دومین تلاش نیز موسوم است. در این راه حل تنها از یک تابلوی وضعیت دو یا چند حالته برای فرآیندها استفاده می‌گردد.
تابلوی وضعیت فرآیند: وضعیت فعلی یک فرآیند را برای فرآیند رقیب به نمایش می‌گذارد. تا اگر یکی از فرآیندها در ناحیه‌ی بحرانی نبود و یا قصد ورود به ناحیه‌ی بحرانی را نیز نداشت یعنی در ناحیه‌ی باقی‌مانده قرار داشت، فرآیند دیگری بتواند به ناحیه‌ی بحرانی دسترسی داشته باشد. بنابراین شرط پیشروی برقرار است.
توجه : تابلوی وضعیت فرآیندها، سبب مستقل شدن فرآیندها از تناوب قطعی و خاصیت الاکلنگی می‌شود. این تابلو، دو وضعیت مختلف را برای یک فرآیند به نمایش، می‌گذارد:
Flag [i] = FALSE: یعنی فرآیند مورد نظر در داخل ناحیه‌ی بحرانی قرار ندارد.
Flag [i] = TRUE: یعنی فرآیند مورد نظر علاقه‌مند است تا در صورت خالی بودن ناحیه‌ی بحرانی، وارد ناحیه‌ی بحرانی گردد. به عبارت دیگر علاقه‌مندی یک فرآیند برای ورود به ناحیه‌ی بحرانی را نشان می‌دهد.

شرط ورود به ناحیه‌ی بحرانی یک فرآیند در این راه حل

شرط لازم

فرآیند، علاقه‌مند به ورود به ناحیه‌ی بحرانی باشد، Flag [فرآیند علاقه‌مند] = TRUE

شرط کافی

فرآیند رقیب در ناحیه‌ی بحرانی نباشد، Flag [فرآیند رقیب] = FALSE

ساختار کلی این راه حل به صورت زیر می‌باشد:

```
Boolean flag [2]= {FALSE, FALSE};
```

```
P۰
P۰(void){
   while (TRUE)
   {
   ①while (flag [1]);/*loop*/
   ②flag [0]= TRUE;
   ③ critical_section ();
   ④flag [0] = FALSE;
    remainder_section ();
   }
 }
```

```
P۱
P۱(void){
   while (TRUE)
   {
   while (flag [0]);/*loop*/
   flag [1]= TRUE;
    critical_section ();
   flag [1] = FALSE;
    remainder_section ();
   }
 }
```

توجه : در این راه حل ابتدا وضعیت فرآیند مقابل بررسی می‌شود و بعد از آن فرآیند، تابلوی وضعیت مربوط به خود را به نشانه‌ی علاقه‌مندی به ورود به ناحیه‌ی بحرانی برابر TRUE مقداردهی می‌کند.

مزایا

رعایت شرط پیشرفت

شرط پیشرفت می‌گفت، فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، در تصمیم‌گیری برای ورود فرآیندهای دیگر به ناحیه‌ی بحرانی شرکت نکند.

فرض کنید فرآیند $P_۱$ در ناحیه‌ی باقی‌مانده‌ی سنگین خود قرار دارد (در حال حاضر در ناحیه‌ی بحرانی قرار ندارد و قصد ورود به ناحیه‌ی بحرانی را هم ندارد.) مانند رسم یک نمودار طولانی، (بعد از خط ④و در این زمان $P_۰$ از ناحیه‌ی بحرانی خود خارج گردد و مجدداً درخواست این ناحیه را داشته باشد، اما فرآیند $P_۱$ همچنان در ناحیه‌ی باقی‌مانده‌ی سنگین خود قرار دارد. از آنجا که فرآیند $P_۱$ پس از خروج از ناحیه‌ی بحرانی و قبل از ورود به ناحیه‌ی باقی‌مانده‌ی سنگین خود، مقدار Flag [1] را برابر FALSE قرار داده است (در خط ④، پس باز هم فرآیند $P_۰$ بدون وجود مانعی وارد ناحیه‌ی بحرانی می‌شود. فرآیند $P_۰$ داخل ناحیه‌ی بحرانی می‌گردد، در حالی که فرآیند $P_۱$ در ناحیه‌ی باقی‌مانده قرار دارد، این یعنی پیشرفت. بنابراین شرط پیشرفت برقرار است.

توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند $P_۰$ برقرار است.

while (flag [1]);
FALSE

معایب

عدم رعایت شرط انحصار متقابل

حالت اول (ورود غیرهمزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی باشد، یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی نباشد. پس در ابتدا Flag [1] = FALSE و Flag [0] = FALSE می‌باشد. فرض کنید، ابتدا پردازنده در اختیار فرآیند P_0 باشد، بنابراین فرآیند P_0 پس از عبور از خط ① به دلیل عدم برقراری شرط حلقه while (Flag [1]); وارد خط ② می‌شود و با مقداردهی Flag [0] = TRUE علاقه‌مندی خود را FALSE برای ورود به ناحیه‌ی بحرانی اعلام می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد.

توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار بود. حال اگر پردازنده به فرآیند P_1 تعویض متن انجام دهد، فرآیند P_1 در حلقه‌ی انتظار مشغول خواهد ماند.

while (flag [0]);
TRUE

تا زمانی که فرآیند P_0 از ناحیه‌ی بحرانی خود خارج شود و Flag [0] = FALSE شود. توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_1 برقرار نبود. زیرا ناحیه‌ی بحرانی پُر است.

بنابراین شرط انحصار متقابل در حالت اول برقرار است.

حالت دوم (ورود تقریباً همزمان فرآیندها)

زمانی که هر دو فرآیند به طور تقریباً همزمان، قصد ورود به ناحیه‌ی بحرانی را دارند. به صورت پیش فرض تابلوی وضعیت هر دو فرآیند هر دو Flag [0] = FALSE و Flag [1] = FALSE می‌باشد.

وقتی که هر دو فرآیند به طور تقریباً همزمان به دستور while می‌رسند خط ① برای فرآیند P_0 while (Flag [0]); و برای فرآیند P_1، whlie (Flag [1]); برقرار نیست.

FALSE FALSE

بنابراین هر دو فرآیند P_0 و P_1 به طور تقریباً همزمان پس از اعلام علاقه‌مندی خود برای ورود به ناحیه‌ی بحرانی با مقداردهی Flag [0] = TRUE برای فرآیند P_0 و Flag [1] = TRUE برای فرآیند P_1، در خط ② در ادامه وارد ناحیه‌ی بحرانی می‌شوند. بنابراین شرط انحصار متقابل در حالت دوم برقرار نیست.

پس در کل، شرط انحصار متقابل برای این راه حل برقرار نیست.

عدم رعایت شرط انتظار محدود

گرسنگی دارد: گرسنگی یعنی اینکه یک فرآیند مدام کار کند و از کار کردن فرآیندهای دیگر جلوگیری کند. در اینجا گرسنگی بدین معنی که یک فرآیند مدام داخل ناحیه‌ی بحرانی برود و از ورود فرآیندی دیگر به ناحیه‌ی بحرانی جلوگیری کند، وجود دارد.

فرض کنید فرآیند P_0 در ناحیه‌ی بحرانی است و با انجام تعویض متن، پردازنده به P_1 داده شود. در این لحظه فرآیند P_1 مشغول بررسی Flag [0] برای کسب اجازه‌ی ورود به ناحیه‌ی بحرانی است که اجازه‌ی ورود به آن داده نمی‌شود و تا پایان کوانتوم خود می‌چرخد، چون مقدار Flag [0] برابر TRUE است. فرض کنید در پایان کوانتوم، با انجام تعویض متن مجدد، پردازنده از فرآیند P_1 گرفته شود و دوباره به فرآیند P_0 داده شود و فرآیند P_0 پس از اجرای ناحیه‌ی بحرانی‌اش، سریعاً ناحیه‌ی باقی‌مانده را پشت سرگذاشته و بخواهد مجدداً وارد ناحیه‌ی بحرانی شود. شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار است، بنابراین فرآیند P_0 مشکلی برای ورود مجدد ندارد، بنابراین مقدار Flag [0] توسط فرآیند P_0، مجدداً TRUE می‌شود. حال اگر در این لحظه، بر اثر تعویض متن پردازنده مجدداً به فرآیند P_1 برسد، از آنجا که Flag [0] قبلاً توسط فرآیند P_0 مقدارش TRUE شده است، پس شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_1 برقرار نیست، لذا به فرآیند P_1 باز هم اجازه‌ی ورود به ناحیه‌ی بحرانی داده نخواهد شد و باز هم P_0 ناحیه‌ی بحرانی را در اختیار دارد. در نتیجه در این راه حل، دستیابی به ناحیه‌ی بحرانی شانسی و تصادفی است و این سناریو می‌تواند بارها و بارها تکرار شود و در نهایت، سبب گرسنگی فرآیند P_1 گردد. بنابراین شرط انتظار محدود به دلیل گرسنگی برقرار نیست.

جهت درک بیشتر، یک بار دیگر به الگوریتم مطرح شده، نگاه کنید، واقعیت این است که هر فرآیند برای ورود به ناحیه‌ی بحرانی ابتدا وضعیت فرآیند مقابل را بررسی می‌کند و بعد از آن، فرآیند، مقدار تابلوی وضعیت مربوط به خود را به نشانه‌ی علاقه‌مندی به ورود به ناحیه‌ی بحرانی برابر TRUE قرار می‌دهد، بنابراین این الگوریتم، ابزاری برای جلوگیری و ممانعت از ورود فرآیندهای رقیب به ناحیه‌ی بحرانی که چابک‌تر هستند، در اختیار ندارد.

مثال واقعی این الگوریتم در عالم زندگی نیز وجود دارد، مانند فردی خجول که برای بدست آوردن عامل مشترک مورد علاقه‌ی خود، ابتدا وضعیت عامل مشترک را بررسی می‌کند و بعد علاقه‌مندی خود را اعلام می‌کند، این فرد در مواجه با افراد چابک‌تر مدام دچار گرسنگی می‌شود! در راه حل قبل ابتدا افراد علاقه‌مندی خود را اعلام می‌کردند و بعد وضعیت عامل مشترک را بررسی می‌کردند، که دیدید بن‌بست را به ارمغان می‌آورد! یکی برای همه به صورت همزمان، این یعنی بن‌بست، دیدید که

هیچ یک از این راه حل‌ها مناسب نبودند، هم باید علاقه‌مندی برای تصاحب عامل مشترک اعلام گردد و هم نوبت رعایت گردد. البته به شرطی که عامل مشترک خالی باشد. **راه حل پترسون همین راگفته بود.**
بن‌بست ندارد: زیرا در ورود تقریباً همزمان فرآیندها بن‌بست وجود نداشت.

راه حل‌های سخت‌افزاری

در این دسته راه حل‌ها، مسئولیت برقراری شرط انحصار متقابل بر عهده‌ی برنامه‌نویس و سخت‌افزار است.

دستورالعمل از کار انداختن وقفه‌ها

ساده‌ترین راه حل، آن است که هر فرآیند بلافاصله پس از ورود به ناحیه‌ی بحرانی‌اش، تمام وقفه‌ها را از کار بیندازد و دقیقاً قبل از خروج از ناحیه‌ی بحرانی همه‌ی آن‌ها را مجدداً فعال سازد. با متوقف کردن وقفه‌ها، دیگر وقفه‌های ساعت نیز رخ نخواهند داد. پردازنده فقط در نتیجه‌ی وقوع وقفه‌های ساعت و یا وقفه‌های ورودی و خروجی است که می‌تواند از یک فرآیند به فرآیندی دیگر تعویض متن کند. پس با از کار انداختن وقفه‌ها، پردازنده دیگر تحت هیچ شرایطی قادر نخواهد بود از فرآیندی به فرآیند دیگر تعویض متن کند. بنابراین هنگامی که یک فرآیند وقفه‌ها را غیرفعال می‌کند، می‌تواند بدون هیچ مشکلی و بدون ترس از مداخله‌ی دیگر فرآیندها به خواندن و نوشتن در حافظه‌ی مشترک بپردازد. پس برای رعایت شرط انحصار متقابل **در سیستم‌های تک پردازنده‌ای** کافی است، وقفه‌ها متوقف شوند.
ساختار کلی این راه حل به صورت زیر می‌باشد:

```
P (int i) {
   while (TRUE)
   {
   Disable_Interupts ();
   critcial_section ();
   Enable_Interupts );
   remainder_section ();
   }
}
```

معایب

۱ـ اگر کاربر وقفه‌ها را خاموش کند ولی بعد از اجرای ناحیه‌ی بحرانی، فراموش کند، مجدداً وقفه‌ها را فعال نماید، سبب از کار افتادن وقفه و در نتیجه کل سیستم می‌شود. پس اعطای قدرت متوقف کردن وقفه‌ها به فرآیند کاربر از نظر معیار امنیت سیستم، عاقلانه نیست.

توجه :از کار انداختن وقفه، در سیستم‌های تک پردازنده‌ای، اغلب در داخل خود سیستم عامل برای مدیریت نواحی بحرانی فرآیندهای سیستم عامل، تکنیک مفیدی است، اما برای مدیریت نواحی بحرانی فرآیندهای کاربر، به دلیل فراموش کار بودن کاربر، تکنیک مناسبی نیست.

۲ـ در سیستم‌های چند پردازنده‌ای، غیرفعال کردن وقفه‌ها، فقط در پردازنده‌ای اثر دارد که دستورالعمل از کار انداختن وقفه را اجرا می‌کند. پس پردازنده‌های دیگر به کارشان ادامه می‌دهند. از آنجایی که ممکن است بیش از یک فرآیند در هر لحظه در حال اجرا باشد، ممکن است فرآیندهای دیگر که روی پردازنده‌های دیگر در حال اجرا هستند نیز وارد ناحیه‌ی بحرانی شوند. به عبارت دیگر، در سیستم‌های چند پردازنده‌ای، این ذات توازی است که باعث ورود همزمان فرآیندها به ناحیه‌ی بحرانی می‌شود و نه وقوع وقفه، که با جلوگیری از آن، بخواهیم مشکل را حل کنیم. پس در سیستم‌های چند پردازنده‌های ممکن است شرط انحصار متقابل برقرار نباشد.

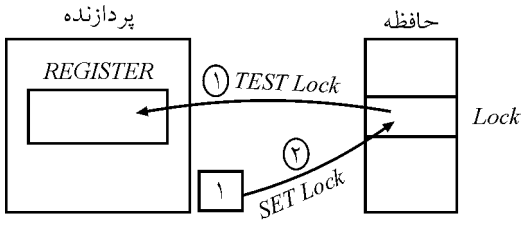
دستورالعمل TSL

راه حل متغیر قفل را به یادآورید، علت عدم موفقیت آن راه حل، عدم قدرت کافی، برای برقراری شرط انحصار متقابل بود و این عدم قدرت، ناشی از اتمیک نبودن (تجزیه‌پذیر بودن) دو دستور خواندن قفل (TEST Lock) و مقداردهی قفل (SET Lock) می‌باشد. در آن راه حل، این امکان وجود داشت که یک فرآیند زمانی که متغیر قفل را می‌خواند و آن را برابر صفر می‌بیند و با ناحیه‌ی بحرانی خالی مواجه می‌شود به جای آنکه فوراً متغیر قفل را به یک مقداردهی کند و سد راه ورود فرآیند رقیب به ناحیه‌ی بحرانی گردد، ممکن است این امکان فراهم شود که پردازنده قبل از آنکه فرآیند فعلی متغیر قفل را به نشانه‌ی تصاحب ناحیه‌ی بحرانی به یک مقداردهی کند، به فرآیند رقیب تعویض متن کند و باعث شود فرآیند رقیب نیز متغیر قفل را بخواند و مقدار آن را برابر صفر ببیند و آن هم با ناحیه‌ی بحرانی خالی مواجه شود و مطابق آنچه پیش از این نیز گفتیم، در ادامه هر دو فرآیند وارد ناحیه‌ی بحرانی شوند.

اگر بتوانیم کاری کنیم که دو عمل خواندن و مقداردهی متغیر قفل به صورت اتمیک (تجزیه‌ناپذیر) انجام شود، مسأله حل می‌شود، دیدید که از کار انداختن وقفه‌ها هم نتوانست یک راه حل عمومی باشد. بسیاری از پردازنده‌ها، دستورالعمل دو بخشی اما اتمیک خاصی دارند به نام TSL (Test and Set Lock)، بدین نحو که این دستورالعمل به شکل تجزیه‌ناپذیر، عملیات زیر را انجام می‌دهد:

- ابتدا محتویات یک کلمه از حافظه به نام Lock را می‌خواند و مقدار آن را در رجیستر قرار می‌دهد. (TEST Lock)
- سپس مقدار یک را در متغیر Lock قرار می‌دهد. (SET Lock)

به شکل زیر توجه کنید:



سخت‌افزار تضمین می‌کند که دو عمل خواندن و مقداردهی متغیر قفل به صورت اتمیک (تجزیه‌ناپذیر) انجام شود. بدین شکل که هیچ فرآیند و حتی پردازنده‌ی دیگری نتواند به این متغیر قفل دسترسی پیدا کند تا وقتی که اجرای دستورالعمل به پایان برسد. پردازنده‌ای که دستورالعمل TSL را اجرا می‌کند، گذرگاه حافظه را قفل می‌کند تا از دسترسی دیگر پردازنده‌ها به حافظه جلوگیری کند تا اینکه این دستورالعمل به پایان برسد.

ساختار کلی این راه حل، به صورت زیر می‌باشد:

```
enter_section ();          enter_section              exit_section
critical_section ();        TSL REGISTER , LOCK        MOVE LOCK , 0
exit_section ();           CMP REGISTER , 0           RET
remainder_section ();      JNE enter-section
                           RET
```

توجه مهم :در این راه حل شرط ورود به ناحیه‌ی بحرانی، اجرای زودتر دستور TSL است.

در این الگوریتم، فرآیندها برای کسب اجازه و ورود به ناحیه‌ی بحرانی از تابع enter_section و برای خروج از ناحیه‌ی بحرانی از تابع exit_section استفاده می‌کنند. اگر فرآیندی علاقه‌مند به ورود به ناحیه‌ی بحرانی باشد. ابتدا تابع enter_section را صدا می‌زند تا بررسی‌های لازم جهت فراهم بودن یا نبودن ورود به ناحیه‌ی بحرانی انجام گردد. بدین نحو که ابتدا، توسط دستور TSL و به شکل اتمیک مقدار متغیر Lock در رجیستر قرار داده می‌شود، سپس مقدار متغیر Lock برابر یک مقداردهی می‌شود. سپس مقدار رجیستر که حاوی مقدار متغیر Lock می‌باشد با صفر مقایسه می‌شود. اگر مقدار رجیستر برابر صفر باشد به معنی خالی بودن ناحیه‌ی بحرانی است و در ادامه دستور JNE به معنی پرش به ابتدای تابع به شرط برابر نبودن مقدار رجیستر با صفر، انجام نمی‌شود، زیرا مقدار رجیستر برابر صفر است. و در ادامه دستور RET (RETURN) باعث می‌شود تا تابع تمام شود. بنابراین فرآیند علاقه‌مند به ورود به ناحیه‌ی بحرانی، وارد ناحیه‌ی بحرانی می‌گردد. و پس از آنکه فرآیند کارش با ناحیه‌ی بحرانی تمام شد، تابع exit_section را به نشانه‌ی خروج از ناحیه‌ی بحرانی صدا می‌زند، اجرای این تابع سبب می‌شود تا مقدار متغیر Lock توسط دستور MOVE برابر صفر گردد، صفر بودن مقدار متغیر Lock به معنی خالی بودن ناحیه‌ی بحرانی است، بنابراین این امکان فراهم می‌شود تا

فرآیندهای دیگر بتوانند پس از کسب اجازه توسط تابع enter-section وارد ناحیه‌ی بحرانی شوند.
در طرف مقابل اگر در حین اجرای تابع enter_section، مقدار متغیر Lock برابر یک باشد، به معنی پُر بودن ناحیه‌ی بحرانی است و در ادامه شرط دستور JNE به معنی پرش به ابتدای تابع به شرط برابر نبودن رجیستر با مقدار صفر برقرار است، زیرا مقدار رجیستر برابر یک است. بنابراین یک حلقه‌ی انتظار مشغول تا خالی شدن ناحیه‌ی بحرانی ایجاد می‌گردد. این حلقه‌ی انتظار مانع ورود فرآیند رقیب به ناحیه‌ی بحرانی می‌گردد.

مزایا

رعایت شرط انحصار متقابل

حالت اول (ورود غیرهمزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی باشد، یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی نباشد. پس در ابتدا مقدار متغیر Lock برابر صفر است.
فرض کنید، فرآیند P_0 با فراخوانی تابع enter_section علاقه‌مندی خود را برای ورود به ناحیه‌ی بحرانی اعلام کند، این تابع دستور TSL (دو عمل خواندن و مقداردهی متغیر قفل) را به صورت اتمیک انجام می‌دهد، بنابراین فرآیند P_0، ناحیه‌ی بحرانی را تصاحب می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد.
توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار بود.
حال اگر پردازنده به فرآیند P_1 تعویض متن کند، این فرآیند نیز تابع enter_section را فراخوانی می‌کند و در دستور CMP مقدار رجیستر را برابر یک می‌بیند و مطابق آنچه پیش از این نیز گفتیم شرایط برای فرآیند P_1 به گونه‌ای رقم می‌خورد که باید در حلقه‌ی انتظار مشغول، مشغول باشد. تا زمانی که فرآیند P_0 از ناحیه‌ی بحرانی خود خارج گردد و تابع exit_section را فراخوانی کند تا مقدار متغیر قفل برابر صفر گردد. (شرایط ورود به ناحیه‌ی بحرانی فراهم گردد) توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_1، برقرار نیست. (ناحیه‌ی بحرانی پُر است). بنابراین شرط انحصار متقابل برقرار است.

حالت دوم (ورود تقریباً همزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی است، یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی نباشد. پس در ابتدا مقدار متغیر Lock برابر صفر است. در این حالت هر دو فرآیند برای کسب اجازه‌ی ورود به ناحیه‌ی بحرانی تابع enter_section را به شکل همزمان فراخوانی می‌کنند. از آن جا که دستور TSL (دو عمل خواندن و مقداردهی قفل) به صورت اتمیک انجام می‌گردد. در نهایت یک فرآیند خوش شانس‌تر که کمی زودتر پردازنده را برای اجرای دستور TSL در اختیار بگیرد، سرانجام ناحیه‌ی

بحرانی را تصاحب می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد و دیگری باید در حلقه‌ی انتظار مشغول، مشغول باشد. بنابراین شرط انحصار متقابل در حالت دوم نیز برقرار است.

رعایت شرط پیشرفت

شرط پیشرفت می‌گفت، فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، در تصمیم‌گیری برای ورود فرآیندهای دیگر به ناحیه‌ی بحرانی نباید شرکت کند در این راه حل یک فرآیند پس از خروج از ناحیه‌ی بحرانی و قبل از ورود به ناحیه‌ی باقی‌مانده، توسط فراخوانی تابع exit_section (مقداردهی متغیر قفل به صفر) ناحیه‌ی بحرانی را خالی اعلام می‌کند، در واقع مانعی که سد راه ورود فرآیندهای دیگر به ناحیه‌ی بحرانی بود، از روی ناحیه‌ی بحرانی برمی‌دارد و در ادامه برای رسیدگی به کارهای دیگرش، در ناحیه‌ی باقی‌مانده قرار می‌گیرد، بنابراین فرآیندهای رقیب می‌توانند پس از کسب اجازه‌ی ورود به ناحیه‌ی بحرانی توسط تابع enter_section و مستقل از فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، وارد ناحیه‌ی بحرانی شوند. بنابراین شرط پیشرفت برقرار است.

معایب

عدم رعایت شرط انتظار محدود

گرسنگی دارد: گرسنگی یعنی اینکه یک فرآیند مدام کار کند و از کار کردن فرآیندهای دیگر جلوگیری کند. در اینجا گرسنگی بدین معنی که یک فرآیند مدام داخل ناحیه‌ی بحرانی برود و از ورود فرآیندهای دیگر به ناحیه‌ی بحرانی جلوگیری کند، وجود دارد. از آن جا که این الگوریتم فاقد مکانیزم نوبت‌دهی است. بنابراین ورود به ناحیه‌ی بحرانی تصادفی خواهد بود، پس یک فرآیند خوش‌شانس‌تر و با بخت و اقبال بالاتر (منظور از شانس، چگونگی تعویض متن پردازنده است) می‌تواند، بارها و بارها پردازنده را بدست آورد و وارد ناحیه‌ی بحرانی گردد، بدون آنکه فرآیند رقیب بتواند کاری از پیش ببرد، این یعنی گرسنگی زیرا سرنوشت ورود فرآیندها به بخت و اقبال فرآیندها گره خورده است، و هیچ گونه نوبتی رعایت نمی‌شود، بنابراین شرط انتظار محدود به دلیل گرسنگی برقرار نیست.

بن‌بست ندارد:

زیرا پس از اجرای موازی یا تقریباً همزمان، هر دو فرآیند برای کسب اجازه‌ی ورود به ناحیه‌ی بحرانی تابع enter_section را به شکل همزمان فراخوانی می‌کنند. از آن جا که دستور TSL (دو عمل خواندن و مقداردهی قفل) به صورت اتمیک انجام می‌گردد. لذا یک فرآیند خوش‌شانس‌تر ناحیه‌ی بحرانی را تصاحب می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد و دیگری باید در حلقه‌ی انتظار مشغول، مشغول باشد. پس هیچگاه هر دو فرآیند به طور همزمان مسدود نمی‌شوند، در نتیجه بن‌بست رخ نخواهد داد، بنابراین شرط انتظار محدود در این شرایط برقرار است.

عدم رعایت انتظار مشغول

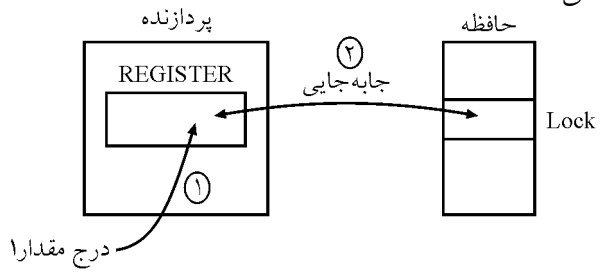
در این راه حل، مشکل انتظار مشغول وجود دارد، زیرا زمانی که مثلاً فرآیند P_0 در ناحیه‌ی بحرانی قرار دارد، فرآیند P_1 در یک حلقه‌ی انتظار زمان پردازنده را در بررسی برقراری شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی به هدر می‌دهد.

توجه: این راه حل، فقط در پردازنده‌هایی قابل اجرا است که دستورالعمل TSL را پشتیبانی می‌کنند.

توجه : پردازنده Intel، دستور TSL را پشتیبانی نمی‌کند و از دستور SWAP پشتیبانی می‌کند.

۳ـ دستورالعمل SWAP

این دستورالعمل محتوای یک رجیستر را با محتوای محلی از حافظه (متغیر Lock) به شکل اتمیک (تجزیه‌ناپذیر) جابه‌جا می‌کند.



سخت‌افزار تضمین می‌کند که عمل جابه‌جایی محتوای رجیستر و حافظه به صورت اتمیک (تجزیه‌ناپذیر) انجام شود. بدین شکل که هیچ فرآیند و حتی پردازنده دیگری نتواند به این متغیر قفل دسترسی پیدا کند تا وقتی که اجرای دستورالعمل به پایان برسد. پردازنده‌ای که دستورالعمل SWAP را اجرا می‌کند، گذرگاه حافظه را قفل می‌کند تا از دسترسی دیگر پردازنده‌ها به حافظه جلوگیری کند تا اینکه این دستورالعمل به پایان برسد.

ساختار کلی این راه حل به صورت زیر می‌باشد:

```
enter_section ();
critical_section ();
exit_section();
remainder_section();
```

```
enter_section
MOVE REGITSTER, 1
SWAP RGISTER, LOCK
CMP REGISTER, 0
JNE enter_section
RET
```

```
exit_section ();
MOVE Lock, 0
RET
```

توجه مهم : در این راه حل شرط ورود به ناحیه‌ی بحرانی، اجرای زودتر دستور swap است.

در این الگوریتم، فرآیندها برای کسب اجازه‌ی ورود به ناحیه‌ی بحرانی از تابع enter_section و برای خروج از ناحیه‌ی بحرانی از تابع exit_section استفاده می‌کنند. اگر فرآیندی علاقه‌مند به ورود به

ناحیه‌ی بحرانی باشد، ابتدا تابع enter_section را صدا می‌زند تا بررسی‌های لازم جهت فراهم بودن یا نبودن ورود به ناحیه‌ی بحرانی انجام گردد. بدین نحو که ابتدا، مقدار رجیستر برابر یک مقداردهی می‌شود. سپس توسط دستور swap و به شکل اتمیک مقدار متغیر Lock و رجیستر جابه‌جا می‌شود، سپس مقدار رجیستر که حاوی مقدار متغیر Lock می‌باشد با صفر مقایسه می‌شود. اگر مقدار رجیستر برابر صفر باشد به معنی خالی بودن ناحیه‌ی بحرانی است و در ادامه دستور JNE به معنی پرش به ابتدای تابع به شرط برابر نبودن مقدار رجیستر با صفر، انجام نمی‌شود، زیرا مقدار رجیستر برابر صفر است و در ادامه دستور RET (RETURN) باعث می‌شود تا تابع تمام شود. بنابراین فرآیند علاقه‌مند به ورود به ناحیه بحرانی وارد ناحیه‌ی بحرانی می‌گردد. و پس از آنکه فرآیند کارش با ناحیه‌ی بحرانی تمام شد، تابع exit_section را به نشانه خروج از ناحیه‌ی بحرانی صدا می‌زند، اجرای این تابع سبب می‌شود تا مقدار متغیر Lock توسط دستور MOVE برابر صفر گردد، صفر بودن مقدار متغیر Lock به معنی خالی بودن ناحیه‌ی بحرانی است، بنابراین این امکان فراهم می‌شود تا فرآیندهای دیگر بتوانند پس از کسب اجازه توسط تابع enter_section وارد ناحیه‌ی بحرانی شوند.

در طرف مقابل اگر در حین اجرای تابع enter_section مقدار متغیر Lock برابر یک باشد، به معنی پُر بودن ناحیه‌ی بحرانی است و در ادامه شرط دستور JNE به معنی پرش به ابتدای تابع به شرط برابر نبودن رجیستر با مقدار صفر، برقرار است، زیرا مقدار رجیستر برابر یک است، بنابراین یک حلقه‌ی انتظار مشغول تا خالی شدن ناحیه‌ی بحرانی ایجاد می‌گردد. این حلقه‌ی انتظار مانع ورود فرآیند رقیب به ناحیه‌ی بحرانی می‌گردد.

مزایا

رعایت شرط انحصار متقابل

حالت اول (ورود غیرهمزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی باشد، یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی نباشد. پس در ابتدا مقدار متغیر Lock برابر صفر است.

فرض کنید، فرآیند P_0 با فراخوانی تابع enter_section علاقه‌مندی خود را برای ورود به ناحیه‌ی بحرانی اعلام کند، این تابع دستور swap (عمل جابه‌جایی مقدار متغیر قفل و رجیستر) را به صورت اتمیک انجام می‌دهد، بنابراین فرآیند P_0 ناحیه‌ی بحرانی را تصاحب می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد.

توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار بود.

حال اگر پردازنده به فرآیند P_1 تعویض متن کند، این فرآیند نیز تابع enter_section را فراخوانی می‌کند و در دستور CMP مقدار رجیستر را برابر یک می‌بیند و مطابق آنچه پیش از این نیز گفتیم شرایط

برای فرآیند P_1 به گونه‌ای رقم می‌خورد که باید در حلقه‌ی انتظار مشغول، مشغول باشد. تا زمانی که فرآیند P_0 از ناحیه‌ی بحرانی خود خارج گردد و تابع exit_section را فراخوانی کند تا مقدار متغیر قفل برابر صفر گردد. (شرایط ورود به ناحیه‌ی بحرانی فراهم گردد) توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_1، برقرار نیست. (ناحیه‌ی بحرانی پُر است). بنابراین شرط انحصار متقابل برقرار است.

حالت دوم (ورود تقریباً همزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی است، یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی نباشد. پس در ابتدا مقدار متغیر Lock برابر صفر است. در این حالت هر دو فرآیند برای کسب اجازه‌ی ورود به ناحیه‌ی بحرانی تابع enter_section را به شکل همزمان فراخوانی می‌کنند. از آن جا که در دستور swap (عمل جابه‌جایی مقدار متغیر قفل و رجیستر) به صورت اتمیک انجام می‌گردد در نهایت یک فرآیند خوش‌شانس‌تر که کمی زودتر پردازنده را برای اجرای دستور swap در اختیار بگیرد، سرانجام ناحیه‌ی بحرانی را تصاحب می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد و دیگری باید در حلقه‌ی انتظار مشغول، مشغول باشد. بنابراین شرط انحصار متقابل در حالت دوم نیز برقرار است.

رعایت شرط پیشرفت

شرط پیشرفت می‌گفت، فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، در تصمیم‌گیری برای ورود فرآیندهای دیگر به ناحیه‌ی بحرانی نباید شرکت کند. در این راه حل یک فرآیند پس از خروج از ناحیه‌ی بحرانی و قبل از ورود به ناحیه‌ی باقی‌مانده، توسط فراخوانی تابع exit_section (مقداردهی متغیر قفل به صفر) ناحیه‌ی بحرانی را خالی اعلام می‌کند، در واقع مانعی که سد راه ورود فرآیندهای دیگر به ناحیه‌ی بحرانی بود، از روی ناحیه‌ی بحرانی برمی دارد و در ادامه برای رسیدگی به کارهای دیگرش، در ناحیه‌ی باقی‌مانده قرار می‌گیرد، بنابراین فرآیندهای رقیب می‌توانند پس از کسب اجازه‌ی ورود به ناحیه‌ی بحرانی توسط تابع enter_section و مستقل از فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، وارد ناحیه‌ی بحرانی شوند. بنابراین شرط پیشرفت برقرار است.

معایب

عدم رعایت شرط انتظار محدود

گرسنگی دارد: گرسنگی یعنی اینکه یک فرآیند مدام کار کند و از کار کردن فرآیندهای دیگر جلوگیری کند. در اینجا گرسنگی بدین معنی که یک فرآیند مدام داخل ناحیه‌ی بحرانی برود و از ورود فرآیندهای دیگر به ناحیه بحران جلوگیری کند، وجود دارد، از آن‌جا که این الگوریتم فاقد مکانیزم نوبت‌دهی است، بنابراین ورود به ناحیه‌ی بحرانی تصادفی خواهد بود، پس یک فرآیند خوش‌شانس‌تر و با بخت

و اقبال بالاتر (منظور از شانس، چگونگی تعویض متن پردازنده است) می‌تواند، بارها و بارها پردازنده را بدست آورد و وارد ناحیه‌ی بحرانی گردد، بدون آنکه فرآیند رقیب بتواند کاری از پیش ببرد، این یعنی گرسنگی زیرا سرنوشت ورود فرآیندها به بخت و اقبال فرآیندها گره خورده است، و هیچ گونه نوبتی رعایت نمی‌شود، بنابراین شرط انتظار محدود به دلیل گرسنگی برقرار نیست.

بن‌بست ندارد:

زیرا پس از اجرای موازی یا تقریباً همزمان، هر دو فرآیند برای کسب اجازه‌ی ورود به ناحیه‌ی بحرانی تابع enter_section را به شکل همزمان فراخوانی می‌کنند. از آن جا که در دستور swap (عمل جابه‌جایی مقدار متغیر قفل و رجیستر) به صورت اتمیک انجام می‌گردد. لذا یک فرآیند خوش‌شانس‌تر ناحیه‌ی بحرانی را تصاحب می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد و دیگری باید در حلقه‌ی انتظار مشغول، مشغول باشد. پس هیچ گاه هر دو فرآیند به طور همزمان مسدود نمی‌شوند، در نتیجه بن‌بست رخ نخواهد داد، بنابراین شرط انتظار محدود در این شرایط برقرار است.

عدم رعایت انتظار مشغول

در این راه حل، مشکل انتظار مشغول وجود دارد، زیرا زمانی که مثلاً فرآیند P_0 در ناحیه‌ی بحرانی قرار دارد، فرآیند P_1 در یک حلقه‌ی انتظار زمان پردازنده را در بررسی برقراری شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی به هدر می‌دهد.

این راه حل، فقط در پردازنده‌هایی قابل اجراست که دستور swap را پشتیبانی می‌کنند.

توجه : پردازنده Intel، از دستور swap، پشتیبانی می‌کند.

راه حل‌های سیستم عامل (سمافور ـ راهنما): Semaphore

از راه‌حل‌های نرم‌افزاری گفته شده الگوریتم پیترسون می‌تواند به عنوان راه حل صحیح به کار رود. همچنین در راه حل‌های سخت‌افزاری گفته شده استفاده از دستورالعمل‌های TSL و swap با صرف نظر کردن از مشکل گرسنگی‌شان، می‌توانند به عنوان یک راه حل استفاده شوند. اما مشکلی که وجود دارد، این است که این نوع راه حل‌ها دو مشکل اساسی زیر را دارند:

۱ ـ انتظار مشغول

همان طور که اشاره شد، راه حل‌های نرم‌افزاری و سخت‌افزاری، یک مشکل اساسی به نام انتظار مشغول دارند. در این دسته راه حل‌ها، زمانی که فرآیندی در ناحیه‌ی بحرانی به سر می‌برد. هر فرآیند دیگری که برای ورود به ناحیه‌ی بحرانی تلاش کند در هر بار تعویض متن پردازنده باید به طور پیوسته داخل یک حلقه، زمان را سپری کند. این حلقه‌های بیهوده یک مشکل محسوب می‌شود. چرا که در زمانی که انتظار مشغول وقت پردازنده را تلف می‌کند، ممکن است فرآیندهای دیگری وجود داشته

باشند که بتوانند از پردازنده استفاده مفید کنند.

۲ ـ اولویت معکوس

فرض کنید دو فرآیند، یکی با اولویت بالا و دیگری با اولویت پایین وجود دارد. اگر فرآیند با اولویت پایین در ناحیه‌ی بحرانی باشد و سپس فرآیند با اولویت بالا وارد حافظه‌ی اصلی شود و درخواست ناحیه‌ی بحرانی را داشته باشد، چون اولویت بالاتری دارد، زمان‌بند، پردازنده را در اختیار آن قرار می‌دهد و منتظر منبع بحرانی می‌ماند. در این حین، فرآیند با اولویت پایین پردازنده را در اختیار ندارد تا کار خود را به پایان رساند و ناحیه‌ی بحرانی را آزاد کند. در نتیجه‌ی این سناریو، فرآیند با اولویت بالا تابی نهایت بار در حلقه‌ی انتظار مشغول، دور می‌زند. لازم به ذکر است که به این وضعیت مشکل اولویت معکوس گفته می‌شود، که همان بن‌بست است.

چون فرآیند با اولویت پایین، منتظر فرآیند با اولویت بالا است تا پردازنده را رها کند و فرآیند با اولویت بالا، منتظر فرآیند با اولویت پایین است تا ناحیه‌ی بحرانی را آزاد کند شرط لازم (انحصار متقابل) برقرار است، شروط بدیهی (انحصاری بودن و نگهداری و انتظار) هم وجود دارد، شرط کافی، (سیکل انتظار چرخشی) هم برقرار است، پس وقوع بن‌بست حتمی است.)

با به کارگیری سمافور به عنوان راه حلی جهت کنترل شرایط رقابتی می‌توان از بروز چنین مشکلاتی جلوگیری کرد. راه حل سمافور در سال ۱۹۶۵ توسط Dijkstra پیشنهاد شده است.

ساختار کلی این راه حل به صورت زیر می‌باشد:

```
wait (mutex);
critical_section ();
signal (mutex);
remainder_section ();
```

تذکر: mutex از عبارت Mutual Exclusion گرفته شده است.

توجه : دو تابع wait (mutex) و signal (mutex)، باید به صورت اتمیک (تجزیه‌ناپذیر) انجام گیرند. اتمیک بودن، تضمین می‌کند که از لحظه‌ای که یک عملیات بر روی شمارنده سمافور شروع می‌شود، هیچ فرآیند دیگری نتواند به سمافور دسترسی پیدا کند تا زمانی که آن عملیات به پایان برسد. اتمیک بودن این عملیات برای حل مسایل همگام‌سازی و کنترل شرایط رقابتی و به تبع برقراری شرط انحصار متقابل کاملاً لازم و ضروری است.

توجه : در مقاله Dijkstra، از نام‌های P و V (حرف اول کلمات آلمانی تست "probern" و افزایش "Verhogen") به ترتیب به جای wait و signal استفاده شده بود و همچنین در سایر متون، از نام‌های down و up به ترتیب برای این دو استفاده می‌کنند. در همان متون گاهاً به جای نام‌های down و up، به

ترتیب از عبارت‌های mutex_lock و mutex_unlock نیز استفاده شده است.
راه حل سمافور بر دو دسته کلی (۱) سمافور عمومی و (۲) سمافور دو دویی می‌باشد، که در ادامه به آن‌ها می‌پردازیم:

سمافور عمومی

سمافور عمومی s از یک شمارنده و یک صف تشکیل شده است.
ساختار کلی سمافور عمومی به صورت زیر است:

```
struct semaphore
{
int count;
Queue Type Queue;
} s;
```

شمارنده‌ی سمافور (s.count)

برای برقراری شرط انحصار متقابل از این شمارنده با مقدار اولیه یک استفاده می‌گردد.

صف سمافور (s.queue)

برای برقراری شرط پیشروی، انتظار محدود، حل مسأله انتظار مشغول و حل مسأله اولویت معکوس از این صف استفاده می‌گردد. فرآیندهای منتظر ورود به ناحیه‌ی بحرانی در این صف نگه‌داری می‌شوند. اگر آزاد شدن یا خروج از این صف به ترتیب ورود باشد، اصطلاحاً به آن **سمافور قوی** می‌گویند و در صورتی که ترتیب خروج مشخص نشده باشد، به آن **سمافور ضعیف** گفته می‌شود. سمافورهای قوی عدم گرسنگی را تضمین می‌کنند، اما در سمافورهای ضعیف این گونه نیست. در این کتاب کلیه سمافورها، از نوع قوی فرض می‌شوند، مگر اینکه نوع سمافور ضعیف بیان شود. سیستم عامل‌ها نیز معمولاً از سمافور قوی استفاده می‌کنند.
بر روی سمافور عمومی s دو تابع wait (s) و signal (s) عملیات ورود و خروج از ناحیه‌ی بحرانی را کنترل می‌کنند.
تابع wait (s): عملیات آن ترتیب شامل، کاهش مقدار شمارنده، تست کردن مقدار شمارنده و احیاناً خواباندن یک فرآیند است.
ساختار این تابع به صورت زیر است:

```
wait (semophore s)
   {
s.count = s.count - 1;
```

```
    if (s.count < 0)
    {
    add this process to s.queue;
    block ();
    }
}
```

توجه : راه حل سمافور و تابع اتمیک wait باید توسط سیستم عامل پشتیبانی گردد، در غیر این صورت می‌توان این راه حل را توسط سرویس‌های سیستم عامل شبیه‌سازی کرد.

شرح تابع: پس از فراخوانی تابع wait(s) توسط یک فرآیند علاقه‌مند به ورود به ناحیه‌ی بحرانی، ابتدا یک واحد از شمارنده سمافور کاسته می‌شود (s.count = s.count - 1)، سپس اگر شرط مربوط به دستور if (s.count < 0) برقرار بود (مقدار شمارنده سمافور منفی بود) این فرآیند داخل صف سمافور قرار گرفته و توسط تابع block مسدود و به خواب می‌رود، یعنی از وضعیت اجرا به وضعیت منتظر منتقل می‌گردد، در غیر این صورت، فرآیند، وارد ناحیه‌ی بحرانی می‌گردد.

توجه : در راه حل‌های قبل از سمافور، در صورتی که یک فرآیند علاقه‌مند به ورود به ناحیه‌ی بحرانی، با ناحیه‌ی بحرانی پُر مواجه می‌شد، پس از پایان هر برش زمانی از وضعیت اجرا به وضعیت آماده منتقل می‌گردید و مجدداً توسط زمان‌بند کوتاه مدت، همان فرآیند این شانس را داشت که مجدداً پردازنده را در اختیار بگیرد و مجدداً شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی را بررسی کند، در حالیکه همچنان ناحیه‌ی بحرانی پُر است که از این مسأله به انتظار مشغول یاد کردیم، همانطور که قبلاً هم گفتیم انتظار مشغول سبب هدر دادن وقت پردازنده می‌گردد، اما در راه حل سمافور در صورتی که یک فرآیند علاقه‌مند به ورود به ناحیه‌ی بحرانی، با ناحیه‌ی بحرانی پُر مواجه شود، در صف سمافور قرار داده می‌شود و حالت فرآیند، از وضعیت اجرا به وضعیت منتظر، تغییر می‌کند. در این شرایط فرآیندی که با ناحیه‌ی بحرانی پُر مواجه شده است، دیگر در صف آماده قرار ندارد و به تبع دیگر این شانس را هم نخواهد داشت تا توسط زمان‌بند کوتاه مدت مجدداً انتخاب گردد تا پردازنده را در اختیار بگیرد و مجدداً شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی را بررسی کند و سبب هدر دادن زمان پردازنده گردد. پس علاوه بر برقراری شروط انحصار متقابل، پیش‌روی و انتظار محدود، مسأله انتظار مشغول و اولویت معکوس نیز توسط راه سمافور حل شد.

توجه : در سمافور عمومی، وقتی شمارنده سمافور، مقدار منفی دارد، قدرمطلق این مقدار، معرف تعداد فرآیندهای بلوکه شده در صف سمافور است.

تابع signal (s) : عملیات آن به ترتیب شامل، افزایش مقدار شمارنده، تست کردن مقدار شمارنده و احیاناً بیدار کردن یک فرآیند است.

ساختار این تابع به صورت زیر است:

```
signal (semaphore s)
{
s.count = s.count + 1
   if (s.count < = 0)
   {
   remove a process from queue;
   wake up ();
   }
}
```

توجه : راه حل سمافور و تابع اتمیک signal باید توسط سیستم عامل پشتیبانی گردد، در غیر این صورت می‌توان این راه حل را توسط سرویس‌های سیستم عامل شبیه‌سازی کرد.

شرح تابع: پس از فراخوانی تابع signal (s) توسط یک فرآیند علاقه‌مند به خروج از ناحیه‌ی بحرانی، ابتدا یک واحد به مقدار شمارنده سمافور اضافه می‌شود (s.count = s.count + 1)، سپس اگر شرط مربوط به دستور if (s.count < = 0) برقرار بود (مقدار شمارنده سمافور مثبت نبود) به معنی وجود فرآیندهای علاقه‌مند ورود به ناحیه‌ی بحرانی که در حال حاضر در صف سمافور قرار دارند، به شکل خروج به ترتیب ورود (FIFO) فقط یک فرآیند به ازای هر بار فراخوانی تابع signal(s) توسط تابع wake up () بیدار شده، یعنی تغییر وضعیت داده و از وضعیت منتظر به صف آماده منتقل می‌گردد. بنابراین این فرآیند پس از حضور در صف آماده‌ی پردازنده، این شانس را دارد تا توسط زمان‌بند کوتاه مدت، انتخاب شود و پردازنده را در اختیار بگیرد و در وضعیت اجرا قرار بگیرد.

به بیان دیگر هر فرآیند که از ناحیه‌ی بحرانی خارج شود، با اجرای تابع signal (s)، فرآیند سر صف سمافور را بیدار می‌کند و اگر صف سمافور خالی باشد و هیچ فرآیند خوابیده‌ای در آن سمافور وجود نداشته باشد در تابع signal فقط یک واحد به مقدار شمارنده سمافور اضافه می‌شود و تابع خاتمه می‌یابد.

مزایا

رعایت شرط انحصار متقابل

حالت اول (ورود غیر همزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی است. یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی قرار ندارد. پس در ابتدا مقدار شمارنده سمافور برابر یک است.

فرض کنید، فرآیند P_0 با فراخوانی تابع wait (s) علاقه‌مندی خود را برای ورود به ناحیه‌ی بحرانی

اعلام کند، سیستم عامل اتمیک (تجزیه‌ناپذیر) بودن اجرای تابع wait (s) را تضمین می‌کند، هنگام اجرای تابع اتمیک wait (s) و پس از اجرای s.count = s.count - 1، مقدار شمارنده سمافور، برابر صفر می‌گردد.

سپس دستور if (s.count < 0) بررسی می‌گردد و چون شرط برقرار نیست تابع wait (s); خاتمه می‌یابد، بنابراین فرآیند P_0 ناحیه‌ی بحرانی را تصاحب می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد. توجه کنید که شروط لازم و کافی برای ورود به ناحیه‌ی بحرانی برای فرآیند P_0 برقرار بود. حال اگر پردازنده به فرآیند P_1 تعویض متن کند، این فرآیند نیز تابع wait (s) را فراخوانی می‌کند و پس از اجرای دستور s.count = s.count -1 مقدار شمارنده سمافور برابر ۱- می‌گردد. سپس دستور if (s.count < 0) بررسی می‌گردد و چون شرط برقرار است فرآیند P_1 شمارنده سمافور به صف سمافور می‌رود. حال اگر پردازنده به فرآیندهای بعدی علاقه‌مند ورود به ناحیه‌ی بحرانی نیز تعویض متن کند، تا زمانی که فرآیند P_0 در ناحیه‌ی بحرانی باشد، وضع به همین منوال خواهد بود.

حالت دوم (ورود تقریباً همزمان فرآیندها)

در شرایطی که ناحیه‌ی بحرانی خالی است، یعنی هیچ فرآیندی داخل ناحیه‌ی بحرانی قرار ندارد. پس در ابتدا مقدار شمارنده سمافور برابر یک است.

در این حالت فرآیندها برای کسب اجازه‌ی ورود به ناحیه‌ی بحرانی تابع wait (s) را به شکل همزمان فراخوانی می‌کنند. از آنجا که تابع wait (s) به صورت اتمیک انجام می‌گردد. یک فرآیند خوش‌شانس‌تر که زودتر پردازنده را بدست می‌آورد تابع wait (s) را زودتر فراخوانی کند، ناحیه‌ی بحرانی را تصاحب می‌کند و در ادامه وارد ناحیه‌ی بحرانی می‌گردد و فرآیندهای بعدی تا زمانی که فرآیند اول داخل ناحیه‌ی بحرانی قرار دارد، به ترتیب در صف سمافور قرار می‌گیرند. بنابراین شرط انحصار متقابل برقرار است.

توجه :اگر مقدار اولیه شمارنده سمافور برابر n می‌بود، همه n فرآیند قادر خواهند بود که همزمان وارد ناحیه‌ی بحرانی شوند و هیچ مشکلی حل نشده است! اگر مقدار اولیه شمارنده سمافور، برابر صفر می‌بود، هر فرآیندی که بخواهد وارد ناحیه‌ی بحرانی شود، بلوکه می‌شود، یعنی پس از مدتی همه فرایندها بلوکه می‌شوند و هیچگاه بیدار نخواهند شد زیرا دیگر کسی نیست که بخواهد آن‌ها را بیدار کند. اما اگر مطابق آنچه بیش از این نیز گفتیم، مقدار شمارنده سمافور را برابر یک در نظر بگیریم، تنها یک فرآیند می‌تواند وارد ناحیه‌ی بحرانی شود و پس از ورود فرآیند اول به ناحیه‌ی بحرانی، مقدار شمارنده سمافور صفر خواهد شد، بنابراین اگر در هنگامی که فرآیند اول در ناحیه‌ی بحرانی به سر می‌برد، یک یا چند فرآیند دیگر قصد ورود به ناحیه‌ی بحرانی را داشته باشند به ترتیب در صف سمافور خواهند خوابید. پس شرط انحصار متقابل برقرار است. در آخر اینکه هر فرآیند که از ناحیه‌ی

بحرانی خارج شود با اجرای تابع signal (s) فرآیند سر صف سمافور را بیدار می‌کند.

رعایت شرط پیشرفت

هرگاه فرآیندی که در ناحیه‌ی باقی‌مانده قرار دارد، در تصمیم‌گیری برای ورود فرآیندهای دیگر به ناحیه‌ی بحرانی شرکت کند، آن راه حل شرط پیشروی را رعایت نمی‌کند. در این راه حل یک فرآیند پس از خروج از ناحیه‌ی بحرانی و قبل از ورود به ناحیه‌ی باقی‌مانده، توسط فراخوانی تابع signal ابتدا ناحیه‌ی بحرانی را خالی اعلام می‌کند، سپس برای رسیدگی به ادامه کارهایش، در ناحیه‌ی باقی‌مانده قرار می‌گیرد. خالی اعلام کردن ناحیه‌ی بحرانی توسط یک فرآیند پس از خروج از ناحیه‌ی بحرانی و قبل از ورود به ناحیه‌ی باقی‌مانده باعث می‌شود تا این فرآیند زمانی که در ناحیه‌ی باقی‌مانده قرار دارد، در تصمیم‌گیری برای ورود فرآیندهای دیگر به ناحیه‌ی بحرانی شرکت نکند و این یعنی برقراری شرط پیشرفت. فراخوانی تابع signal توسط یک فرآیند خارج شده از ناحیه‌ی بحرانی باعث می‌شود تا فرآیند ابتدای صف سمافور از خواب بیدار شود و از وضعیت منتظر به وضعیت آماده تغییر حالت دهد و در صف آماده پردازنده قرار گیرد. در واقع یک فرآیند پس از خروج از ناحیه‌ی بحرانی توسط فراخوانی تابع signal راه را برای ورود فرآیند ابتدای صف سمافور به صف آماده پردازنده فراهم می‌کند تا در صورتی که زمان‌بند کوتاه مدت، پردازنده را در اختیارش قرار داد، بتواند ناحیه‌ی بحرانی را تصاحب کند.

این روند یکی پس از دیگری و به صورت خروج به ترتیب ورود برای صف سمافور رخ می‌دهد، هر فرآیندی که از ناحیه‌ی بحرانی خارج شود، فرایند ابتدای صف سمافور را بیدار می‌کند و این روال ادامه پیدا می‌کند، بنابراین شرط پیشرفت برقرار است.

رعایت شرط انتظار محدود

گرسنگی ندارد: گرسنگی یعنی اینکه یک فرآیند مدام کار کند و از کار کردن فرآیندهای دیگر جلوگیری کند. در اینجا گرسنگی بدین معنی که یک فرآیند مدام داخل ناحیه‌ی بحرانی برود و از ورود فرآیندهای دیگر به ناحیه‌ی بحرانی جلوگیری کند، وجود ندارد. از آن جا که فرآیندهای بلوکه شده در صفی از فرآیندها که در آن ترتیب خروج به ترتیب ورود (FIFO) است، قرار می‌گیرند، پس هیچگاه گرسنگی بروز نمی‌کند.

بن‌بست ندارد: استفاده از سمافور مطابق آنچه در حالت دوم برقراری شرط انحصار متقابل گفتیم، ایجاد بن‌بست نمی‌کند. زیرا ورود به ناحیه‌ی بحرانی نوبتی انجام می‌شود.

توجه : استفاده نادرست از سمافور ممکن است، منجر به بن‌بست گردد.

مثال: دو فرآیند P_0 و P_1 را در نظر بگیرید که هر کدام به دو سمافور s و q با مقادیر اولیه یک برای

شمارنده سمافور دسترسی دارند، کد فرآیندها به صورت زیر است:

P_0	P_1
① wait (s);	② wait (q);
③ wait (q);	④ wait (s);
⋮ ⋮	
⑤ signal (s);	⑥ signal (q);
⑦ signal (q);	⑧ signal (s);

فرض کنید دستورات ①و ②با هم (در سیستم چند پردازنده‌ای و اجرای موازی) و یا یک در میان (در سیستم تک پردازنده‌ای و اجرای همروند) اجرا شوند. حال اجرای دستور ③باید منتظر اجرای دستور ⑥باشد و اجرای دستور ④ منتظر اجرای دستور ⑤باشد. این بدان معناست که بن‌بست رخ داده است.

توجه :در برخی کتب غیر مرجع، راه حل سمافور را دارای مشکل بن‌بست می‌دانند، حتی به غلط نتیجه گرفته‌اند که در راه حل سمافور شرط انتظار محدود به دلیل بن‌بست رعایت نمی‌شود. اما دیدید که کنترل شرایط رقابتی برای برقراری شروط انحصار متقابل، پیشروی و انتظار محدود به سادگی با یک سمافور بدون هیچ نقصی امکان‌پذیر است.

در واقع باید بگوییم که تنها مشکل سمافور در سیستمی که حافظه مشترک دارد، پیچیدگی استفاده از آن در حل مسایل است. اگر برنامه نویس دقت نکند، راه حل سمافور ممکن است مانند مثال قبل، دچار بن‌بست شود. در این‌صورت، نباید بگوییم سمافور مشکل بن بست دارد! یعنی ترجیح می‌دهیم این مشکل را به هوش برنامه‌نویس نسبت دهیم. این مشکل بر این نکته اشاره دارد که باید در هنگام استفاده از سمافورها دقیق باشید. یک خطای کوچک و ظریف ممکن است، همه چیز را متوقف کند.

سمافور دودویی

کلیه تعاریف، توضیحات و خصوصیاتی که برای سمافور عمومی گفته شد، برای سمافور دودویی نیز برقرار است. تنها تفاوت سمافور دودویی و سمافور عمومی در نحوه‌ی تعریف توابع wait و singal است.

تابع wati(s) :عملیات آن به ترتیب شامل، تست کردن مقدار شمارنده، کاهش مقدار شمارنده (اما در نهایت فقط تا مقدار صفر) و احیاناً خواباندن یک فرآیند.

ساختار این تابع به صورت زیر است:

```
wait (semaphore s)
```

```
{
    if (s.count == 1)
    s.count = 0
    else
    {
    add this process to s.queue;
    block ();
    }
}
```

توجه :به تفاوت ترتیب عملیات سمافور عمومی و دودویی در تابع wait دقت کنید.

تابع signal (s):عملیات آن به ترتیب شامل تست خالی بودن صف، افزایش مقدار شمارنده و احیاناً بیدار کردن یک فرآیند است.

```
signal (semaphore s)
  {
    if (s.queue is empty)
      s.count = 1;
      else
      {
      remove a process from queue;
      wake up ();
      }
}
```

توجه :به تفاوت ترتیب عملیات سمافور عمومی و دودویی در تابع signal دقت کنید.

توجه : در سمافور دودویی، شمارنده سمافور، هیچگاه منفی نمی‌شود و در هر شرایطی فقط می‌تواند دو مقدار باینری یا دودویی ۰ یا ۱ را داشته باشد، برای اثبات، مجدداً به تعاریف توابع wait و signal در سمافور دودویی دقت کنید.

توجه : هر فرآیندی که از ناحیه‌ی بحرانی خارج شود با اجرای تابع signal (s)، فرآیند ابتدای صف را بیدار می‌کند (تغییری در شمارنده سمافور ایجاد نشده و برابر مقدار صفر باقی می‌ماند) و اگر هیچ فرآیند منتظری درون صف سمافور وجود نداشته باشد، مقدار شمارنده سمافور به یک مقداردهی می‌شود.

توجه مهم : در حل مسائل، سمافور عمومی و دودویی، یک پاسخ یکسان را تولید خواهند کرد، اما شیوه‌های رسیدن به پاسخ با توجه به متفاوت بودن توابع wait و signal، کمی تفاوت دارد، ما در این کتاب برای حل کلیه مسائل سمافور، از سمافور عمومی استفاده می‌کنیم، به شما نیز توصیه می‌کنیم، همین راه حل را در پیش بگیرید.

همگام سازی فرآیندها توسط سمافور

سمافور دو کاربرد دارد (۱) کنترل شرایط رقابتی (۲) همگام سازی فرآیندها
کاربرد سمافور در کنترل شرایط مسابقه بررسی شد.
حال با یک مثال به بیان کاربرد سمافور در همگام‌سازی فرآیندها می‌پردازیم.

مثال: دو فرآیند P_1 و P_2 را در نظر بگیرید که می‌خواهند به طور همروند اجرا گردند. P_1 دستور S_1 و P_2 دستور S_2 را اجرا می‌کند. فرض کنید دستور S_2 فقط باید بعد از دستور S_1 اجرا شود. برای این کار یک شمارنده سمافور مشترک به نام synch و با مقدار صفر را برای P_1 و P_2 در نظر گرفته و کدهای زیر را برای P_1 و P_2 می‌نویسیم.

P_1	P_2
S_1;	wait (synch);
signal (synch);	S_2;

در این پیاده‌سازی، پُر واضح است که تا زمانی که S_1 از فرآیند اول اجرا نشود، فرآیند P_2 اجازه اجرای خط S_2 را ندارد. چون مقدار اولیه synch صفر است، P_2 فقط هنگامی S_2 را اجرا می‌کند که P_1 قبلاً signal(synch) را صدا زده باشد و این کار نیز فقط بعد از دستور S_1 انجام می‌پذیرد.

نکته مهم: برای همگام سازی، معمولاً مقدار اولیه شمارنده سمافور برابر صفر و برای برقراری شرط انحصار متقابل مطابق آنچه پیش از این نیز گفتیم همیشه برابر یک است.
در ادامه حل مسائل کلاسیک با استفاده از راه حل سمافور مورد بررسی قرار می‌گیرد.

حل مسأله تولیدکننده و مصرف‌کننده توسط سمافور (با بافر محدود)

دو فرآیند که یکی تولیدکننده و دیگری مصرف‌کننده است، یک بافر با اندازه‌ی محدود را به اشتراک می‌گذارند. فرآیند تولیدکننده اطلاعات را در بافر قرار می‌دهد و فرآیند مصرف‌کننده اطلاعات را از بافر برمی‌دارد. مشکل زمانی به وجود می‌آید که تولیدکننده می‌خواهد اطلاعاتی را در بافر قرار دهد اما با بافر پُر مواجه می‌شود که در این حالت تولیدکننده باید بخوابد تا مصرف‌کننده توسط عمل برداشت اطلاعات، مقداری از بافر را خالی کند، تا شرایط بیدار شدن تولیدکننده فراهم گردد. به همین ترتیب اگر مصرف‌کننده بخواهد داده‌ای را از بافر بردارد ولی با بافر خالی مواجه شود، باید بخوابد تا تولیدکننده

توسط عمل درج اطلاعات، مقداری از بافر را پُر کند، تا شرایط بیدار شدن مصرف‌کننده فراهم گردد. در این راه حل از سه شمارنده سمافور استفاده می‌شود، اولی که full نامیده شده است برای شمردن تعداد خانه‌های پُر بافر و دومی که empty نامیده شده است برای شمارش خانه‌های خالی بافر و سومی هم که mutex نامگذاری شده است برای برقراری شرط انحصار متقابل به کار می‌رود، mutex به این منظور استفاده می‌شود که مطمئن شویم تولیدکننده و مصرف‌کننده به طور همزمان به بافر دسترسی ندارند. در ابتدا مقدار اولیه full=۰ و empty = N یعنی برابر تعداد کل خانه‌های بافر و mutex=۱ می‌باشد. شمارنده‌های سمافوری که مقدار اولیه آنها برابر یک است اغلب برای این استفاده می‌شوند که از ورود همزمان دو یا چند فرآیند به ناحیه‌ی بحرانی جلوگیری کنند. (برقراری شرط انحصار متقابل) اگر هر فرآیند، تابع wait را قبل از ورود به ناحیه‌ی بحرانی و تابع signal را دقیقاً پس از خروج از ناحیه‌ی بحرانی اجرا کند، شرط انحصار متقابل برقرار خواهد شد.

ساختار کلی این راه حل برای حل مسأله تولیدکننده و مصرف‌کننده به صورت زیر می‌باشد:

```
تولیدکننده (producde)
while (TRUE)
{
تولید قطعه ;
wait (empty);
wait (mutex);
درج قطعه تولید شده در بافر (ناحیه‌ی بحرانی)
signal (mutex);
signal (full);
}
```

```
مصرف‌کننده (consumer)
while (TRUE)
{
wait (full);
wait (mutex);
حذف قطعه از بافر (ناحیه‌ی بحرانی)
signal (mutex);
signal (empty);
مصرف قطعه حذف شده ;
}
```

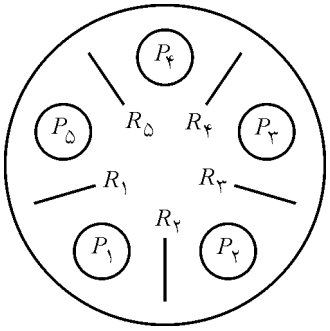
در قطعه کدهای فوق شمارنده‌های سمافور به دو روش و دو کاربرد کاملاً متفاوت استفاده شده‌اند، شمارنده سمافور mutex جهت برقراری شرط انحصار متقابل استفاده شده است تا تضمین کند در هر لحظه فقط یک فرآیند از بافر استفاده می‌کند (یا تولیدکننده یا مصرف‌کننده). ولی سمافورهای full و empty جهت همگام‌سازی فرآیندهای تولید کننده و مصرف کننده استفاده شده‌اند تا تضمین کنند، اتفاقات با ترتیب مشخص انجام می‌پذیرد. در این حالت، مطمئن می‌شویم که تولیدکننده در هنگامی که بافر پُر است، متوقف می‌شود و مصرف‌کننده نیز در هنگامی که بافر خالی است، متوقف می‌شود. این کاربرد کاملاً با انحصار متقابل متفاوت است.

حل مسأله تولید کننده و مصرف کننده توسط سمافور (با بافر نامحدود)

جهت حل مسأله تولیدکننده و مصرف‌کننده با بافر نامحدود توسط سمافور، کافی است شمارنده سمافور empty و به تبع توابع wait (empty) و signal (empty) را از راه حل فوق حذف کنید. در این شرایط تولید کننده، بدون مانع wait (empty)، تا هر چه قدر که می‌خواهد، می‌تواند اقدام به تولید قطعه نماید!

حل مسأله فیلسوفان خورنده توسط سمافور

۵ فیلسوف زندگی خود را صرف فکر کردن و خوردن کرده‌اند. آنها دور یک میز دایره‌ای با ۵ بشقاب و ۵ عدد چنگال نشسته‌اند. هر فیلسوف برای غذا خوردن حتماً باید دو چنگال در دست داشته باشد. بین هر جفت بشقاب روی میز، یک چنگال وجود دارد. هنگامی که فیلسوفی در حال فکر کردن است با بقیه هیچ ارتباطی ندارد. هر از گاهی فیلسوف احساس گرسنگی کرده و سعی می‌کند، چنگال‌های سمت راست و چپش را یکی یکی و با هر ترتیب ممکن بردارد، اگر موفق به برداشتن هر دو چنگال شود برای مدتی غذا خورده و دوباره چنگال‌ها را پایین گذاشته و به فکر کردن ادامه می‌دهد. فیلسوف مجاز است که در هر بار فقط یک چنگال را بردارد و همچنین نمی‌تواند چنگالی که دست فیلسوف دیگری است را به زور بگیرد.



راه حل اول

از آنجایی که هر فیلسوفی که در حال خوردن است، نمی‌تواند چنگال‌های در اختیارش را به فیلسوف مجاور بدهد، لازم است برای هر چنگال شرط انحصار متقابل برقرار باشد. لذا برای هر چنگال یک شمارنده سمافور تعریف می‌شود. مسأله فیلسوفان خورنده در راه حل اول بر دو نوع (۱) چپگرد و (۲) راستگرد می‌باشد.

راه حل چپگرد: فیلسوفان ابتدا چنگال سمت چپ را بر می‌دارند.

ساختار کلی این راه حل به صورت زیر می‌باشد:

تذکر: قطعه کد زیر را شبه کد فرض کنید.

```
# define N 5
semaphor fork[5] ={1};
void philosopher (int i)
{
   while (TRUE)
      {
         wait (fork [i]);
         wait (fork [(i+ 1) % N]);
         eat ();
         signal (fork [i]);
         signal (fork [(i+ 1) % N]);
         think ();
         }
}
```

برای هر چنگال (R_i) ، یک شمارنده سمافور با مقدار اولیه یک تعریف شده است، مطابق الگوریتم فوق هر فیلسوفی که می‌خواهد تغذیه کند، باید بتواند ابتدا چنگال چپ خود را بردارد، این کار با دستور wait (fork [i]) انجام می‌گیرد. سپس با اجرای دستور wait (fork [(i+ 1) % N]) باید بتواند چنگال راست خود را بردارد. اگر فیلسوفی موفق به انجام این دو عمل شد، می‌تواند خوردن را شروع کند.

اگر چه این راه حل تضمین می‌کند که هیچ دو همسایه‌ای همزمان غذا نخورند، ولی این روش ممکن است دچار بن‌بست شود.

فرض کنید که هر پنج فیلسوف همزمان گرسنه شده و هر کدام چنگال سمت چپ خود را بردارند (همه‌ی ۵ فیلسوف خط اول یعنی wait (fork [i]);) را اجرا کنند). بنابراین تمام عناصر آرایه fork[5] برابر با صفر می‌شوند. هنگامی که فیلسوفان سعی می‌کنند تا چنگال سمت راست خود را بردارند wait(fork[(i+ 1)%5);، یک بن‌بست به وجود می‌آید.

شرایط بن‌بست را به یادآورید:

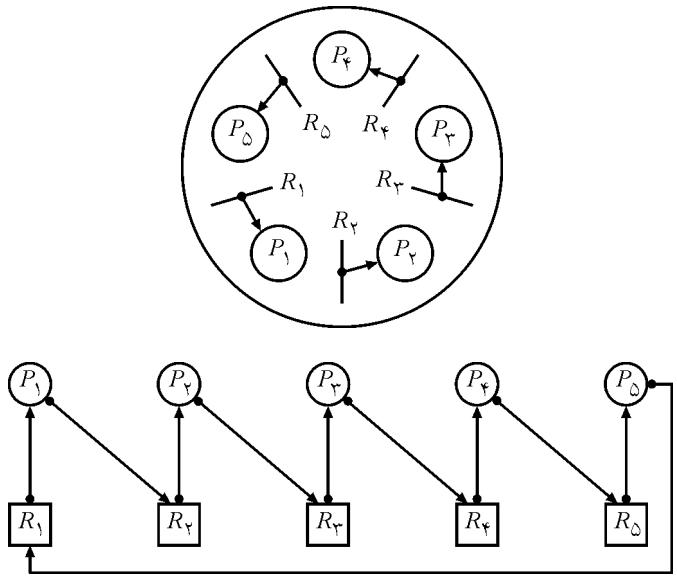
۱ـ انحصار متقابل (برقرار است، توسط تعریف شمارنده سمافور برای هر چنگال (منبع بحرانی))

۲ـ انحصاری بودن (برقرار است، نمی‌توان چنگال (منبع بحرانی) را به زور پس گرفت)

۳ـ نگهداری و انتظار (برقرار است)

۴ـ سیکل انتظار چرخشی (برقرار است)

به شکل بازسازی شده‌ی فیلسوفان خورنده در شرایط بن‌بست توجه کنید:



راه حل راستگرد: فیلسوفان ابتدا چنگال سمت راست را برمی‌دارند.

ساختار کلی این راه حل به صورت زیر می‌باشد:

```
# define N 5
semaphore fork[5] = {1};
void philsopher (int i)
   {
      while (TRUE)
         {
         wait (fork [(i+ 1)%N]);
         wait (fork [i]);
         eat();
         signal (fork[(i+ 1)%N]);
         signal (fork[i]);
         think();
         }
}
```

توجه :کلیه تعارف، توضیحات و خصوصیاتی که برای راه حل چپگرد گفته شد، به شکل بالعکس برای راه حل راستگرد نیز برقرار است.

راه حل دوم: در این راه حل هر فیلسوف تنها زمانی می‌تواند تغذیه کند که هر دو چنگال موجود باشند. در واقع یک راه حل بهبود یافته نسبت به راه حل اول که منجر به بن‌بست نشود. این است که پنج با دستورالعمل قبل از ;()think را با یک شمارنده سمافور mutex با مقدار اولیه یک حفاظت کنیم.

ساختار کلی این راه حل به صورت زیر می‌باشد:

تذکر: قطعه کد زیر را شبه کد فرض کنید.

```
# define N 5
semaphore fork [5] = {1}
semaphore mutex =1;
void philsopher (int i)
{
while (TRUE)
   {
      wait (mutex);
      wait (fork [i]);
      wait (fork [(i+ 1) % N]);
      eat ();
      signal (fork [i]);
      signal (fork [(i+ 1) % N ])
      signal (mutex);
      think ();
   }
}
```

هر فیلسوف قبل از شروع به برداشتن چنگال‌ها تابع wait را بر روی یک شمارنده سمافور mutex انجام می‌دهد و بعد از پایین گذاشتن چنگال‌ها تابع signal را بر روی شمارنده سمافور mutex انجام می‌دهد. از نقطه نظر تئوری این راه حل مناسب است، اما کارایی را کاهش می‌دهد، زیرا در هر لحظه فقط یک فیلسوف می‌تواند مشغول غذا خوردن باشد. اما با ۵ چنگال موجود باید بتوانیم به دو فیلسوف اجازه غذا خوردن همزمان را بدهیم.

راه حل مانیتور

اگرچه راه حل سمافور، راه کار درستی برای کنترل شرایط رقابتی و همگام‌سازی می‌باشد ولی استفاده از آن دقت بسیار زیادی را می‌خواهد. مثلاً اگر برنامه‌نویس در نحوه‌ی استفاده و یا چیدمان توابع wait و signal حول ناحیه‌ی بحرانی برای برقراری شرط انحصار متقابل دقت نکند ممکن است شرط انحصار متقابل نقض گردد و یا پدیده بن‌بست آشکار گردد. همچنین اگر برنامه‌نویس در چیدمان این توابع برای همگام‌سازی فرآیندها دقت نکند، ممکن است باز هم پدیده‌ی بن‌بست ایجاد گردد و یا روالی غیر از آنچه مد نظر بوده است انجام گیرد.

برای اینکه نوشتن برنامه‌های درست، ساده‌تر شود، یک ابزار راحت‌تر برای کنترل شرایط رقابتی و همگام‌سازی به نام مانیتور ابداع شد. مانیتور یک راه حل مبتنی بر کامپایلر زبان برنامه‌نویسی است. توابع wait و signal در راه حل سمافور اغلب توسط سیستم عامل پشتیبانی می‌شود، حتی می‌توان این توابع را توسط زبان‌های برنامه‌نویسی شبیه سازی کرد و در کتابخانه توابع قرار داد تا هرگاه نیاز بود، مورد استفاده قرار گیرد. اما مانیتور ماهیتی دارد که باید توسط کامپایلر زبان برنامه‌نویسی پشتیبانی گردد.

ساختار کلی این راه حل به صورت زیر می‌باشد:

```
monitor mon_name
Begin
i: integer;              variable declarations
c: contdition;           condition variable declarations
function f1()
begin
critical_section1 ();
end;
functionf2()
begin
critical_section2 ();
end;
function fn()
begin
critical_sectionn ();
end;
initialization code;
end monitor
```

کاربردهای مانیتور

مانیتور در دو کاربرد متفاوت می‌تواند مورد استفاده قرار گیرد:

۱ـ کنترل شرایط رقابتی (برقراری شرط انحصار متقابل، پیشرفت و انتظار محدود)

۲ـ همگام‌سازی (کنترل شرایط رقابتی + همگام سازی)

در ادامه به تشریح این دو مورد می‌پردازیم:

کنترل شرایط رقابتی با استفاده از مانیتور (برقراری شرط انحصار متقابل)

مانیتور برای کنترل شرایط رقابتی شامل مجموعه‌ای از متغیرهای داده‌ای و توابع می‌باشد که در یک ماژول بسته‌بندی شده است. متغیرهای داده‌ای تعریف شده در داخل مانیتور، فقط توسط توابع داخل همان مانیتور قابل دسترسی هستند. بنابراین فرآیندها برای دسترسی به متغیرهای داده‌ای داخل یک مانیتور، می‌بایست توابع داخل همان مانیتور را فراخوانی کنند.

مهمترین قانون مانیتور برای کنترل شرایط رقابتی، این است که هیچ دو فرآیندی نمی‌توانند به طور همزمان وارد مانیتور شوند، به بیان دیگر در هر لحظه فقط یک فرآیند می‌تواند داخل یک مانیتور فعال باشد. اگر یک فرآیند با فراخوانی یک تابع مانیتور، وارد آن مانیتور شود، هیچ فرآیند دیگری نمی‌تواند با فراخوانی همان تابع یا توابع دیگر، وارد آن مانیتور شود، مگر آنکه فرآیند اول با اتمام اجرای تابع، از مانیتور خارج گردد. (مانیتور خالی گردد) و یا درون مانیتور غیرفعال گردد (جلوتر شرح داده می‌شود) اینکه کامپایلر چگونه شرط انحصار متقابل را برقرار می‌کند، از نظر برنامه‌نویس ضروری نمی‌باشد. برنامه‌نویس فقط کافی است بداند با قرار دادن ناحیه‌ی بحرانی مورد نظر خود در داخل یکی از توابع مانیتور، هیچ وقت دو یا چند فرآیند، همزمان وارد این ناحیه‌ی بحرانی نخواهد شد. چرا که اگر فرآیند اولی یک تابع مانیتور را فراخوانی کرده باشد، اگر فرآیند دومی قبل از پایان یافتن کار فرآیند اول با تابع مانیتور، بخواهد همان تابع فراخوانی شده توسط فرآیند اول و یا حتی تابع دیگری از مانیتور را فراخوانی کند، اجازه‌ی چنین کاری به او داده نخواهد شد و این فرآیند دوم در وضعیت مسدود در صف مانیتور قرار داده می‌شود، به عبارت دیگر هنگامی که یکی از فرآیندها یک تابع مانیتور را اجرا می‌کند، اگر پردازنده بر اثر تعویض متن در اختیار فرآیند دیگری قرار گیرد و آن فرآیند تابع در اختیار فرآیند اول و یا حتی تابع از مانیتور دیگری از مانیتور را فراخوانی کند، فرآیند دوم به شکل مسدود در صف مانیتور قرار می‌گیرد.

هنگامی که یک فرآیند پس از اتمام کار با تابع موردنظر مانیتور، از مانیتور خارج می‌گردد، ساختار مانیتور به شکلی ایجاد شده است که باعث می‌شود یکی از فرآیندهای موجود در صف مانیتور به شکل خروج به ترتیب ورود از وضعیت مسدود به آماده منتقل گردد تا شرایط قرارگیری آن فرآیند در صف آماده پردازنده فراهم گردد.

کامپایلر یک زبان برنامه‌نویسی که مانیتور را پشتیبانی می‌کند، در ابتدا و انتهای توابع مانیتور تعدادی دستورالعمل کنترلی قرار می‌دهد. این دستورالعمل‌های کنترلی که از دید برنامه‌نویس پنهان هستند، این امکان را فراهم می‌کنند تا هر وقت یک فرآیند، یک تابع مانیتور را فراخوانی نمود، ابتدا بررسی شود که آیا اکنون فرآیند دیگری داخل مانیتور قرار دارد یا خیر. اگر فرآیندی داخل مانیتور قرار داشت، فرآیند فراخوانی کننده، مسدود و در داخل صف ورود به مانیتور قرار داده می‌شود، در غیر اینصورت وارد مانیتور می‌گردد. همچنین هرگاه اجرای یک تابع مانیتور متعلق به فرآیند فراخوانی کننده پایان یافت، دستورالعمل‌های پایانی و پنهان موجود در انتهای تابع مانیتور باعث می‌گردد، یک فرآیند از صف مانیتور به شکل خروج به ترتیب ورود از وضعیت مسدود به آماده منتقل گردد تا شرایط برقراری آن فرآیند در صف آماده پردازنده فراهم گردد.

کامپایلر برای پیاده‌سازی مانیتور و تحقق انحصار متقابل و نیز صف‌بندی فرآیندهای منتظر ورود به مانیتور، می‌تواند در ورودی مانیتور از سمافور mutex استفاده کند. مثلاً در ابتدای همه‌ی توابع مانیتور، یک تابع wait (mutex) و در انتهای همه توابع مانیتور، یک تابع signal (mutex) قرار دهد. بدیهی است این نحوه‌ی پیاده‌سازی و برقراری شرط انحصار متقابل درون مانیتور، توسط سمافور از دید برنامه‌نویس سطح کاربر پنهان است و توسط کامپایلر انجام می‌گردد.

توجه: شمارنده سمافور mutex در این حالت، برای همه توابع مانیتور به صورت مشترک مورد استفاده قرار می‌گیرد.

مانیتور مانند یک ساختمان چند طبقه می‌باشد، که در هر طبقه یک عامل مشترک خاص قرار داده شده است. مثلاً در طبقه‌ی اول استخر و سونا، در طبقه‌ی دوم امکانات ورزشی، در طبقه‌ی سوم سینما و در طبقات بعدی هم عامل‌های مشترک خاص دیگری قرار داده شده است. اما مطابق قوانین این ساختمان، در هر لحظه فقط یک نفر می‌تواند داخل این ساختمان حضور داشته باشد. تا با خیالی آسوده و راحت از عامل مشترک موجود در یک طبقه‌ی خاص استفاده کند. مثلاً اگر فرد اولی در طبقه‌ی سوم و در حال استفاده از عامل مشترک سینما باشد و افراد دیگری علاقه‌مندی خود را برای ورود به طبقه‌ی سوم و یا حتی طبقات دیگر اعلام کنند، نگهبان ساختمان جلوی این افراد را می‌گیرد و آن‌ها را به ترتیب به صف ورودی درب ساختمان هدایت می‌کند. تا وقتی که فرد اول از طبقه‌ی سوم و به تبع از ساختمان خارج گردد. نگهبان ساختمان پس از مشاهده‌ی خروج این فرد، یک نفر را به شکل خروج به ترتیب ورود از صف ورودی درب ساختمان انتخاب و به ساختمان راه می‌دهد!

مثال: کنترل شرایط رقابتی در مانیتور

دو فرآیند هم روند P_1 و P_2 در یک سیستم اشتراک زمانی که از متغیر مشترک سراسری s در بخشی از کد خود استفاده می‌کنند، در نظر بگیرید، شرط انحصار متقابل را توسط مانیتور حل کنید؛ (متغیر s

داخل ناحیه‌ی بحرانی استفاده می‌گردد)

P_1	P_2
⋮	⋮
critical_section	critical_section
⋮	⋮

برای حل این مسأله کافی است، نواحی بحرانی دو فرآیند داخل توابع مانیتور قرار گیرند:

```
Monitor cs
s: integer;
procedure prc1
begin
critcal_section;
end;
procedure prc2
begin
critical_section
end;
s: = 0;
end Monitor
```

حال استفاده از نواحی بحرانی فرآیندهای P_1 و P_2، پس از قرار دادن نواحی بحرانی داخل توابع مانیتور به شکل زیر بازنویسی می‌شود:

P_1	P_2
⋮	⋮
cs.prc1	cs.prc2
⋮	⋮

توجه: در واقع نواحی بحرانی مربوط به فرآیندهای P_1 و P_2 داخل توابع prc1 و prc2 موجود در مانیتور قرار گرفتند و مطابق قوانین مانیتور هیچ دو فرآیندی نمی‌توانند به طور همزمان وارد مانیتور شوند، به بیان دیگر در هر لحظه فقط یک فرآیند می‌تواند داخل یک مانیتور فعال باشد. بنابراین شرط انحصار متقابل برقرار است.

همگام سازی با استفاده از مانیتور (کنترل شرایط رقابتی + همگام‌سازی)

مانیتور برای کنترل شرایط رقابتی توام با همگام‌سازی، شامل مجموعه‌ای از متغیرهای داده‌ای، متغیرهای شرطی و توابع می‌باشد که در یک ماژول بسته‌بندی شده است. در برخی مسائل مانند تولیدکننده و مصرف‌کننده علاوه بر نیاز به کنترل شرایط رقابتی مطابق آنچه پیش از این در مورد مانیتور گفتیم، ما به راه حلی نیاز داریم که وقتی فرآیندها نمی‌توانند پیشروی کنند، داخل خود مانیتور مسدود شوند، در مسأله تولیدکننده و مصرف کننده قرار دادن تست‌هایی برای تشخیص پُر یا خالی بودن بافر در داخل تابع مانیتور ساده است. اما چگونه باید یک تابع را هنگامی که متوجه می‌شود بافر پُر است، مسدود کرد؟ راه حل را باید در متغیرهای شرطی مانیتور جستجو کرد. بنابراین ویژگی دیگر مانیتور این است که علاوه بر تعریف متغیرهای داده‌ای مثل، integer، می‌توان متغیرهای شرطی از نوع conditon تعریف کرد، به این‌ها متغیرهای شرطی (condition variables) گفته می‌شود.

بر روی متغیر شرطی x یک مانیتور، دو تابع خاص wait (x) و signal (x) عملیات همگام‌سازی بین دو فرآیند را کنترل می‌کنند:

تابع wait (x):

عملیات آن به ترتیب شامل وارد کردن یک فرآیند به صف متغیر شرطی مانیتور و خواباندن (تغییر وضعیت از اجرا به مسدود) همان فرآیند است.

ساختار این تابع به صورت زیر است:

```
wait (condition x)
{
   add this process to x.queue;
   block ();
}
```

توجه: به تفاوت تابع wait در سمافور و مانیتور دقت کنید، در تابع wait مانیتور، هیچ تغییری، بر روی مقدار متغیر شرطی x اعمال نمی‌گردد و فقط عمل درج یک فرآیند، داخل صف متغیر شرطی و خواباندن همان فرآیند انجام می‌گردد.

شرح تابع: فراخوانی تابع wait (x) متعلق به متغیر شرطی x، توسط یک تابع مانیتور مورد استفاده‌ی یک فرآیند موجود در مانیتور باعث می‌شود تا فرآیندی که در حال حاضر داخل مانیتور قرار دارد، داخل صف متغیر شرطی مانیتور قرار گرفته و توسط تابع ()blockr مسدود شود، یعنی از وضعیت اجرا به وضعیت مسدود منتقل گردد. دقت کنید که فرآیند مسدود شده است اما همچنان داخل مانیتور قرار دارد، اگر فرآیندی توسط تابع wait متعلق به یک متغیر شرطی داخل یک مانیتور مسدود گردد، در این

شرایط فرآیند رقیب می‌تواند وارد مانیتور گردد، زیرا فرآیند اول داخل مانیتور قرار دارد، اما فعال نیست! (خوابیده است). بنابراین ممکن است در یک لحظه بیش از یک فرآیند داخل مانیتور قرار بگیرد، اما فقط یکی از آن‌ها فعال باشد. یکی خواب و دیگری بیدار!

تابع signal (x):

عملیات آن به ترتیب شامل خارج کردن یک فرآیند به شکل خروج به ترتیب ورود از صف متغیر شرطی مانیتور و بیدار کردن (تغییر وضعیت از مسدود به آماده) همان فرآیند است.

ساختار این تابع به صورت زیر است:

```
siganl (condition x)
  {
    remove a process from queue;
    wakeup ();
  }
```

توجه : به تفاوت تابع signal در سمافور و مانیتور دقت کنید، در تابع signal مانیتور، هیچ تغییری، بروی مقدار متغیر شرطی x اعمال نمی‌گردد و فقط عمل حذف یک فرآیند، از صف متغیر شرطی موردنظر به شکل خروج به ترتیب ورود و بیدار کردن همان فرآیند، انجام می‌گردد.

شرح تابع: فراخوانی تابع signal(x) متعلق به متغیر شرطی x، توسط یک تابع مانیتور مورد استفاده‌ی یک فرآیند موجود در مانیتور باعث می‌شود یک فرآیند به شکل خروج به ترتیب ورود از صف متغیر شرطی مانیتور خارج شده و توسط تابع ()wakeup بیدار شود، یعنی از وضعیت منتظر به وضعیت آماده منتقل گردد.

نکته : هنگامی که یک فرآیند از ابتدای صف یک متغیر شرطی توسط تابع signal متعلق به یک متغیر شرطی، بیدار و فعال می‌گردد، قادر خواهد بود تا ادامه دستورات تابع مانیتور را که قبلاً به دلیل مسدود شدن انجام نشده بود، ادامه دهد و تابع مانیتور را تمام کند و از مانیتور خارج گردد.

توجه : تابع signal (x) متعلق به متغیر شرطی x، در صورت استفاده در یک تابع مانیتور باید آخرین دستور تابع مانیتور باشد. زیرا بر اثر اجرای تابع signal، یک فرآیند از ابتدای صف متغیر شرطی که قبلاً توسط تابع wait (x) متعلق به متغیر شرطی x مسدود شده بود، فعال می‌گردد. و اگر بعد از تابع signal باز هم دستور باشد، بدین ترتیب در یک زمان دو فرآیند فعال در مانیتور وجود خواهد داشت و این برخلاف قوانین مانیتور است. بنابراین برای برقراری قوانین داخلی مانیتور، تابع signal در صورت استفاده در توابع مانیتور باید آخرین دستور تابع مانیتور باشد.

در ادامه حل مسأله کلاسیک تولیدکننده و مصرف‌کننده با استفاده از راه حل مانیتور مورد بررسی قرار

می‌گیرد.

حل مسأله تولیدکننده و مصرف کننده توسط مانیتور (با بافر محدود)

```
Monitor ProducerConsumer;
count: integer;
full , empty: condition;
procedure insert (item: integer)
begin
   if count = N then wait (full);
   insert_item (item) درج قطعه تولید شده در بافر (ناحیه‌ی بحرانی)
   count : = count + 1 ;
   if count = 1 then signal (empty);
end;
function remove: integer
begin
   if count =  then wait (empty);
   remove: = remove_item;
   count: = count - 1;
   if count = N - 1 then signal (full)
end;
count : = 0 ;
end monitor

procedure producer
begin
   while true do
   begin
   item: = produce - item; تولید قطعه
   ProducerConsumer.insert (item) فراخوانی تابع درج مانیتور، برای درج قطعه تولید شده
   end
```

```
end;
Procedure consumer
begin
while true do
   begin
   item: = ProducerConsumer.remove; فراخوانی تابع حذف مانیتور برای حذف قطعه
   consume_item (item) مصرف قطعه
   end
End.
```

فرآیند تولیدکننده

وقتی فرآیند تولیدکننده توسط تابع insert مانیتور متوجه می‌شود که به دلیل پُر شدن بافر دیگر قادر به ادامه نیست، تابع wait (full) را بر روی متغیر شرطی full فراخوانی می‌کند، این کار باعث می‌شود، فرآیند تولیدکننده در صف متغیر شرطی full قرار گیرد و در داخل مانیتور مسدود گردد، در این لحظه به دلیل اینکه هیچ فرآیندی در داخل مانیتور فعال نیست، به فرآیند مصرف کننده که قبلاً درخواست ورود به مانیتور را داشته و در صف ورود به مانیتور به شکل مسدود قرار گرفته است. اجازه‌ی ورود به داخل مانیتور داده می‌شود. فرآیند مصرف‌کننده می‌تواند از طریق اجرای تابع signal بـر روی هـمان متغیر شرطی full که فرآیند تولیدکننده روی آن wait کرده بود، او را بیدار سازد.

توجه :انحصار متقابل خودکار در توابع مانیتور تضمین می‌کند که اگر مثلاً تـولیدکننده در درون تـابع insert متوجه شود که بافر پُر است، قادر خواهد بود عملیات wait را کامل کند و بخوابد و نگران این نباشد که مبادا زمان‌بند دقیقاً قبل از تکمیل wait به فرآیند مصرف‌کننده سوئیچ کند، زیرا در این صورت فرآیند مصرف کننده فعلاً اجازه‌ی ورود به مانیتور را پیدا نخواهد کرد و باید صبر کند تا تولیدکننده، فراخوان wait را تمام کند و بخوابد.

فرآیند مصرف کننده

به همین ترتیب وقتی فرآیند مصرف کننده توسط تابع remove مانیتور متوجه می‌شود که به دلیل خالی بودن بافر دیگر قادر به ادامه نیست، تابع wait (empty) را بر روی متغیر شرطی empty فراخوانـی می‌کند، این کار باعث می‌شود، فرآیند مصرف‌کننده در صف متغیر شرطی empty قرارگیرد و در داخل مانیتور مسدود گردد، در این لحظه به دلیل اینکه هیچ فرآیندی در داخل مانیتور فعال نیست، به فرآیند تولیدکننده که قبلاً درخواست ورود به مانیتور را داشته و در صف ورود به مانیتور به شکل مسدود قرار گرفته است، اجازه‌ی ورود به داخل مانیتور داده می‌شود. فرآیند تولیدکننده می‌تواند از طریق اجرای

تابع signal بر روی همان متغیر شرطی empty که فرآیند مصرف کننده روی آن wait کرده بود، او را بیدار سازد.

توجه :انحصار متقابل خودکار در توابع مانیتور تضمین می‌کند که اگر مثلاً مصرف کننده در درون تابع remove متوجه شود که بافر خالی است، قادر خواهد بود عملیات wait را کامل کند و بخوابد و نگران این نباشد که مبادا زمان‌بند دقیقاً قبل از تکمیل wait به فرآیند تولیدکننده سوئیچ کند، زیرا در این صورت فرآیند تولیدکننده فعلاً اجازه‌ی ورود به مانیتور را پیدا نخواهد کرد و باید صبر کند تا مصرف‌کننده فراخوان wait را تمام کند و بخوابد.

راه حل تبادل پیام

هدف از شبکه‌های کامپیوتری تبادل داده میان یک مبدأ و یک مقصد مورد نظر است. با قرار دادن داده (پیام) و آدرس مبدأ و مقصد داخل یک بسته، می‌توان این بسته را به سمت مقصد ارسال و هدایت نمود. بنابراین پیام یک مکانیزم ساده و مناسب جهت تبادل داده و همگام‌سازی و ارتباط بین فرآیندهاست که قابل استفاده در سیستم‌های تک پردازنده‌ای و چند پردازنده‌ای و همچنین سیستم‌های توزیع شده می‌باشد. ارتباط بین فرآیندها توسط دو تابع سیستمی send و receive انجام می‌گردد.

تابع (پیام و آدرس مقصد) send: ارسال پیام به مقصد مورد نظر

تابع (پیام و آدرس مبداء) receive: دریافت پیام از یک مبداء مورد نظر

توجه : یک پیام برای رسیدن به مقصد مورد نظر نیاز به سه آدرس PORT و IP و MAC دارد.

همگام سازی فرستنده و گیرنده

حالت‌های فرآیند فرستنده

حالت همگام: فرآیند فرستنده پس از ارسال پیام تا آمدن پیام تصدیق (ACK) مبنی بر درست رسیدن پیام مسدود می‌گردد. زیرا ممکن است پیام با خطا به مقصد برسد، یا اصلاً به مقصد نرسد، بنابراین پیام دوباره باید ارسال گردد، حال اگر فرآیند فرستنده پس از ارسال مسدود نشود، ممکن است فرآیند فرستنده، پیام بعدی را بر روی پیام فعلی روی نویسی کند. در حالی که شاید ارسال دوباره پیام فعلی لازم باشد.

حالت ناهمگام: فرآیند فرستنده، به شکل چند نخی می‌باشد و هر نخ وظیفه‌ی خاصی را بر عهده دارد. در این حالت نخی که وظیفه‌ی ارسال را بر عهده دارد، پس از ارسال پیام تا آمدن پیام تصدیق (ACK) مبنی بر درست رسیدن پیام مسدود می‌گردد. اما سایر نخ‌ها به وظایف خود می‌پردازند، مثلاً نخی که آماده‌سازی پیام بعدی را بر عهده دارد، به انجام وظایف خود می‌پردازد.

حالت‌های فرآیند گیرنده

حالت همگام: فرآیند گیرنده پس از بررسی بافر و مشاهده‌ی خالی بودن آن، تا دریافت پیام از سوی فرستنده، مسدود می‌گردد.

حالت ناهمگام: فرآیند گیرنده، به شکل چند نخی می‌باشد و هر نخ وظیفه‌ی خاصی را بر عهده دارد. در این حالت نخی که وظیفه‌ی دریافت را بر عهده دارد، تا دریافت پیام از سوی فرستنده، مسدود می‌گردد. اما سایر نخ‌ها به وظایف خود می‌پردازند.

قابلیت اطمینان

در حین تبادل داده ما بین فرستنده و گیرنده، ممکن است شبکه پیام‌ها را گم کند، برای حفاظت در مقابل گم شدن پیام‌ها، فرستنده و گیرنده می‌توانند با یکدیگر توافق کنند که به محض رسیدن پیام، گیرنده یک ACK (پیام تصدیق) مخصوص بفرستد (Acknowledgement). اگر فرستنده در یک فاصله زمانی مشخص و از قبل تعیین شده، ACK را دریافت نکرد. یعنی time out شد، پیام را دوباره ارسال می‌کند. در این حالت پیام اصلاً به مقصد نرسیده است که بخواهد ACK تولید شود حال فرض کنید که پیام درست برسد ولی ACK در راه بازگشت گم شود.

بنابراین باز هم time out می‌شود و فرستنده پیام را دوباره ارسال خواهد کرد. بنابراین گیرنده دوباره آن را دریافت می‌کند. فرض کنید دو بار پیام برداشت از حساب شما اجرا گردد، چه می‌شود؟ بنابراین این مسأله بسیار ضروری است که گیرنده بتواند یک پیام جدید را از ارسال مجدد یک پیام قدیمی تشخیص دهد. معمولاً این مسأله با قرار دادن شماره ترتیب در پیام‌ها حل می‌شود. اگر گیرنده یک پیام دریافت کند که شماره ترتیب آن مشابه شماره قبلی باشد، متوجه می‌شود که این پیام تکراری است و از آن چشم‌پوشی می‌کند.

لازم به ذکر است که پیام ACK که از سوی گیرنده به فرستنده ارسال می‌شود حامل دو پیام مهم است: (۱) درست رسیدن پیام فعلی (۲) اعلام شماره ترتیب پیام بعدی.

مثلاً اگر گیرنده یک پیام با شماره ترتیب ۱، را درست دریافت کند، در پیام ACK که به سمت فرستنده ارسال می‌کند، شماره ترتیب ۲ را از فرستنده درخواست می‌کند تا بفرستد. حال اگر این ACK گم شود، فرستنده همان پیام شماره ترتیب ۱ را دوباره ارسال می‌کند، اما گیرنده منتظر دریافت پیام شماره ۲ است، بنابراین متوجه می‌شود این پیام شماره با ترتیب ۱ تکراری است و آن را دور می‌اندازد و مجدداً ACK را ارسال می‌کند و به فرستنده می‌گوید منتظر پیامی با شماره ترتیب ۲ است.